

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 18:51:00 ; Search time 76.16 Seconds
(without alignments)
40.626 Million cell updates/sec

Title: US-10-773-274A-1
Perfect score: 41
Sequence: 1 ASTTNYT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	8	1 AAP83007	Aap83007 Blocking
2	41	100.0	8	1 AAP91830	Aap91830 Octapepti
3	41	100.0	8	1 AAP91832	Aap91832 Octapepti
4	41	100.0	8	2 AAR13201	Aar13201 Sialic ac
5	41	100.0	8	2 AAR26900	Aar26900 Herpes si
6	41	100.0	8	2 AAR28939	Aar28939 Peptide T
7	41	100.0	8	2 AAR44336	Aar44336 Peptide T
8	41	100.0	8	2 AAR44334	Aar44334 Peptide T
9	41	100.0	8	2 AAR44335	Aar44335 Peptide T
10	41	100.0	8	2 AAR44342	Aar44342 Peptide T
11	41	100.0	8	2 AAR47827	Aar47827 Peptide T
12	41	100.0	8	2 AAR59486	Aar59486 Prototype
13	41	100.0	8	2 AAR59487	Aar59487 Peptide T
14	41	100.0	8	2 AAR59493	Aar59493 Peptide T
15	41	100.0	8	2 AAR61488	Aar61488 Peptide T
16	41	100.0	8	2 AAR61482	Aar61482 Peptide T
17	41	100.0	8	2 AAR61481	Aar61481 Prototype
18	41	100.0	8	2 AAW45929	Aaw45929 Complex g
19	41	100.0	8	2 AAW54934	Aaw54934 HIV gp120
20	41	100.0	8	2 AY30971	Ay30971 Non-Cross
21	41	100.0	8	3 AY67951	Ay67951 Multiple
22	41	100.0	8	4 AAB92347	Aab92347 Virus rel
23	41	100.0	8	4 AAB92346	Aab92346 Virus rel
24	41	100.0	8	4 AAB92331	Aab92331 Virus rel
25	41	100.0	8	4 AAB92330	Aab92330 Virus rel

ALIGNMENTS

RESULT 1

AAP83007
ID AAP83007 standard; protein; 8 AA.

XX AC AAP83007;

XX DT 25-MAR-2003 (revised)

XX DT 10-DEC-1990 (first entry)

XX DE Blocking peptide used in composition for treatment and diagnosis of HIV infections.

XX KW HIV; gp110; monoclonal antibodies; neutralising region.

XX OS Homo sapiens.

XX PN GB2196634-A.

XX PD 05-MAY-1988.

XX PF 19-AUG-1987; 87GB-00019587.

XX PR 20-AUG-1986; 86US-00898273.

XX PR 01-MAY-1987; 87US-00045026.

XX PR 29-JUN-1987; 87US-00067996.

XX PR 07-OCT-1987; 87US-00105761.

XX PA (GENE-) GENETIC SYSTEMS CORP.

XX PI Shriver MK;

XX WPI; 1988-103268/15.

XX PT Monoclonal antibodies and peptide(s) - useful for treatment and diagnosis of human immunodeficiency virus infections.

XX Claim 41; Page 24; 25pp; English.

XX CC A blocking peptide contg. at least 5 adjacent amino acids from this HIV gp110 envelope glycoprotein sequence inhibits virus proliferation. It is useful, in conjunction with a monoclonal antibody to a neutralising region of HIV, for treatment or prophylaxis of HIV infections. N- and C-terminals can have up to 20 amino acids attached. See also AAP83003-06 and AAP83008-12. This is equivalent to NL8701950 (accession no. is from this). (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct DR field.)

XX SQ Sequence 8 AA;

Aab99510 Immediate
Aab99509 Immediate
Aab62037 CTL activ
Aab62038 CTL activ
Adn14264 HIV gp120
Aaw05584 Ependymin
Aar98995 Peptide T
Aap71548 AIDS viru
Aar28918 Peptide f
Aar47826 Peptide T
Aap52855 HIV-1 gp1
Aam52331 HIV SF2-g
Aap60181 Sequence
Aar29711 env prote
Aaw33613 ARV-2 env
Aay77307 HIV-1 (AT
Aae35787 ARV-2 env
Aab57558 AIDS asse
Aab63191 ARV-2 env
Aar28959 HIV gp160

26 41 100.0 8 4 AAB99510
27 41 100.0 8 4 AAB99509
28 41 100.0 8 6 ABR62037
29 41 100.0 8 6 ABR62038
30 41 100.0 8 7 ADN14264
31 41 100.0 9 2 AAW05584
32 41 100.0 9 2 AAR98995
33 41 100.0 10 1 AAP71548
34 41 100.0 10 2 AAR28918
35 41 100.0 10 2 AAR47826
36 41 100.0 71 5 ABP52855
37 41 100.0 451 5 AAM52331
38 41 100.0 474 1 AAP60181
39 41 100.0 474 2 AAR29711
40 41 100.0 474 2 AAW33613
41 41 100.0 474 3 AAY77307
42 41 100.0 474 5 AAE35787
43 41 100.0 474 6 AAB57558
44 41 100.0 474 6 AAB63191
45 41 100.0 509 2 AAR28959

Mon Mar 21 08:14:28 2005

us-10-773-274a-1.rag

Query Match 100.0%; Score 41; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ASTTTNYT 8
 |||||
 Db 1 ASTTTNYT 8

RESULT 2
 AAP91830
 ID AAP91830 standard; peptide; 8 AA.
 XX
 AC AAP91830;
 XX
 DT 25-MAR-2003 (revised)
 DT 10-MAR-2003 (revised)
 DT 17-DEC-2001 (revised)
 DT 22-MAY-1990 (first entry)
 XX
 DE Octapeptide which inhibits HIV binding to cell surface receptors.
 XX
 KW HIV; psoriasis; neuropsychiatric disorders.
 XX
 OS Unidentified.
 XX
 PN USN7352313-N.
 XX
 PD 10-OCT-1989.
 XX
 PF 16-MAY-1989; 89US-00352313.
 XX
 PR 27-MAY-1988; 88US-00199873.
 PR 16-DEC-1988; 88US-00285559.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PA (USDC) US SEC OF COMMERCE.
 XX
 PI Bridge P, Goodwin FK;
 XX
 DR WPI; 1989-378088/51.
 XX
 XX Use of short peptide(s) to inhibit binding of HIV to human cells - for
 PT treating psoriasis and neuro-psychiatric disorder including memory
 PT deficiency and mood disorder.
 XX
 PS Disclosure; Page 8; 18pp; English.

CC This peptide is used in a composition for treatment of chronic conditions
 CC eg neuropsychiatric disorders and psoriasis through inhibition of HIV-
 CC cell surface receptor (CD4) binding. CD4 site is common to the CNS and
 CC immune system. AIDS- and non AIDS-related psoriasis both respond
 CC favourably to treatment with the peptide-containing composition. There
 CC are a range of forms which the composition and administration can take.
 CC The amino-terminal Alanine is D-Ala. (Note: Revised entry submitted to
 CC correct the patent number format of US Government-owned NTIS applications
 CC to prevent clashes with ongoing US granted patent numbers. For further
 CC information please visit the Derwent web site at
 CC www.derwent.com/dwpi/updates/ntis.us.html.) (Updated on 10-MAR-2003 to
 CC add missing OS field.) (Updated on 25-MAR-2003 to correct PF field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 8 AA;
 Query Match 100.0%; Score 41; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ASTTTNYT 8
 |||||
 Db 1 ASTTTNYT 8

RESULT 3
 AAP91832
 ID AAP91832 standard; protein; 8 AA.
 XX
 AC AAP91832;
 XX
 DT 25-MAR-2003 (revised)
 DT 10-MAR-2003 (revised)
 DT 17-DEC-2001 (revised)
 DT 22-MAY-1990 (first entry)
 XX
 DE Octapeptide which inhibits HIV binding to cell surface receptors.
 XX
 KW HIV; psoriasis; neuropsychiatric disorders.
 XX
 OS Unidentified.
 XX
 PN USN7352313-N.
 XX
 PD 10-OCT-1989.
 XX
 PF 16-MAY-1989; 89US-00352313.
 XX
 PR 27-MAY-1988; 88US-00199873.
 PR 16-DEC-1988; 88US-00285559.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PA (USDC) US SEC OF COMMERCE.
 XX
 PI Bridge P, Goodwin FK;
 XX
 DR WPI; 1989-378088/51.
 XX
 XX Use of short peptide(s) to inhibit binding of HIV to human cells - for
 PT treating psoriasis and neuro-psychiatric disorder including memory
 PT deficiency and mood disorder.
 XX
 PS Disclosure; Page 8; 18pp; English.

CC This peptide is used in a composition for treatment of chronic conditions
 CC eg neuropsychiatric disorders and psoriasis through inhibition of HIV-
 CC cell surface receptor (CD4) binding. CD4 site is common to the CNS and
 CC immune system. AIDS- and non AIDS-related psoriasis both respond
 CC favourably to treatment with the peptide-containing composition. There
 CC are a range of forms which the composition and administration can take.
 CC The amino-terminal Alanine is D-Ala and the carboxy-terminal has an amide
 CC gp. attached to Threonine. (Note: Revised entry submitted to correct the
 CC patent number format of US Government-owned NTIS applications to prevent
 CC clashes with ongoing US granted patent numbers. For further information
 CC please visit the Derwent web site at
 CC www.derwent.com/dwpi/updates/ntis.us.html.) (Updated on 10-MAR-2003 to
 CC add missing OS field.) (Updated on 25-MAR-2003 to correct PF field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 8 AA;
 Query Match 100.0%; Score 41; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ASTTTNYT 8
 |||||
 Db 1 ASTTTNYT 8

RESULT 4
 AARI3201
 ID AARI3201 standard; protein; 8 AA.
 XX
 AC AARI3201;
 XX
 DT 25-MAR-2003 (revised)

RESULT 4
 AARI3201
 ID AARI3201 standard; protein; 8 AA.
 XX
 AC AARI3201;
 XX
 DT 25-MAR-2003 (revised)

DT 10-OCT-1991 (first entry)
 XX Sialic acid-bonded polypeptide (3).
 DE Sialic acid; cataract; immune disorder.
 KW Synthetic.
 XX
 OS Key Location/Qualifiers
 FH Modified-site 1
 FT /note= "N-terminally glycosylated by 5-acetamido-
 FT 2,4,7,8,9-penta-O-acetyl-3,5-deoxy-beta- D-glycero-D-
 FT galactononulopyranosonyl"
 XX
 XX JP03151398-A.
 PN 27-JUN-1991.
 XX
 PD 06-NOV-1989; 89JP-00288560.
 XX
 PF 06-NOV-1989; 89JP-00288560.
 XX
 PR (MECT-) MECT CORP.
 XX
 PA WPI; 1991-233839/32.
 XX
 DR New sialic acid derivs. bonded to physiologically active polypeptide -
 XX for treatment of cataracts, immune disorders etc. with prolonged half-
 PT life.
 XX
 PS Example 5; Page 6; 7pp; Japanese.
 XX
 SS The prod. has prolonged half-life and is used as a pharmaceutical for
 CC treatment of various diseases, such as cataract and immune disorders. It
 CC comprises a peptide, N-terminally glycosylated by (opt. acetylated)
 CC sialic acid. See also AAR12932, AAR13162 and AAR13201. (Updated on 25-MAR
 CC -2003 to correct PA field.)
 XX
 SQ Sequence 8 AA;
 Query Match 100.0%; Score 41; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASTTTNYT 8
 DB 1 ASTTTNYT 8
 RESULT 5
 AAR26900
 ID AAR26900 standard; peptide; 8 AA.
 XX
 AC AAR26900;
 XX
 DT 23-SEP-2004 (revised)
 DT 25-MAR-2003 (revised)
 DT 20-MAY-1998 (first entry)
 XX
 XX Herpes simplex virus treatment peptide.
 DE HSV; HSV-1; HSV-2; Varicella Zoster virus; human cytomegalovirus;
 KW Epstein-Barr virus.
 KW
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 1
 FT /label= D-Ala
 FT Modified-site 8
 FT /label= Thr-NH2
 XX
 PN W09214751-A1.

XX 03-SEP-1992.
 PD
 XX 24-FEB-1992; 92WO-DK000053.
 PF
 XX 25-FEB-1991; 91DK-00000319.
 PR
 XX (CARL-) CARLBIO TECH LTD AS.
 PA
 XX Pedersen O, Macfadden DK;
 PI
 XX WPI; 1992-316120/38.
 DR
 XX New linear or cyclic peptide(s) including cysteine residue - occupy
 PT lymphocyte CD4 receptors, for treating and preventing virus infections
 PT with Herpes simplex, Varicella zoster, cytomegalovirus and Epstein-Barr
 PT virus.
 XX
 PS Claim 5; Page 15; 21pp; English.
 XX
 SS The peptide is useful in the treatment and prevention of diseases or
 CC conditions caused by Herpes viruses, esp. Herpes Simplex Virus-1 (HSV-1),
 CC HSV-2, Varicella Zoster Virus (VZV), human cytomegalovirus (HCMV) or
 CC Epstein-Barr Virus (EBV). It can also be used to treat diseases and
 CC conditions caused by the family of herpes viruses in animals, e.g.
 CC Aujeszky's disease in pigs, bovine rhinotracheitis, rhinopharyngitis in
 CC horses, laryngotracheitis in poultry and Marek's disease in chickens. It
 CC can also be used in diagnostics. The peptide has never been reported to
 CC penetrate any type of cells, but only to occupy the CD4 receptors found
 CC on lymphocytes and it apparently penetrates the skin with relative ease.
 CC A major advantage of the peptide is its almost complete lack of toxicity,
 CC which means that the use of high doses for an extended period of time is
 CC possible without any drawbacks. It is an example of a highly generic
 CC peptide (AAR26905,6). See also AAR26901-R26906. (Updated on 25-MAR-2003
 CC to correct PN field.) (Updated on 25-MAR-2003 to correct PF field.)
 XX
 XX Revised record issued on 23-SEP-2004 : Correction to Feature Table Key
 SQ Sequence 8 AA;
 Query Match 100.0%; Score 41; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASTTTNYT 8
 DB 1 ASTTTNYT 8
 RESULT 6
 AAR28939
 ID AAR28939 standard; peptide; 8 AA.
 XX
 AC AAR28939;
 XX
 DT 25-MAR-2003 (revised)
 DT 23-MAR-1993 (first entry)
 XX
 XX Peptide T.
 DE
 XX CFS; fatigue; tension; anger; confusion; peptide T; HIV; synthetic;
 KW human immunodeficiency virus; gp 120; protein kinase A; enhancer.
 KW
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 1
 FT /note= "Ala, D-Ala"
 FT Modified-site 8
 FT /note= "may be amidated"
 XX
 PN W09219257-A1.

PD 12-NOV-1992.
 XX
 PF 08-MAY-1992; 92WO-US003582.
 XX
 PR 08-MAY-1991; 91US-00696556.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 XX
 PI Bridge TP, Goodwin FK;
 XX
 DR WPI; 1992-398525/48.
 XX
 PT Peptide compns. for treating chronic fatigue syndrome - ameliorate
 FT symptoms and improve vigour and cognitive and neuro-motor performance.
 XX
 PS Disclosure; Page 18; 2ipp; English.
 XX
 CC Peptide T is a protein kinase A enhancer which produces both symptomatic
 CC and functional improvement in patients with chronic fatigue syndrome, and
 CC has a sequence derived from a sub region of HIV gp120 responsible for
 CC binding to brain membrane and human T cells. The sequence may be used to
 CC design a synthetic peptide which may be used for treatment of chronic
 CC fatigue syndrome not associated with HIV infection. The peptide will
 CC reduce fatigue, tension, anger and confusion and improves cognitive and
 CC neuromotor performance. See also AAR28918-42. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX
 SQ Sequence 8 AA;
 Query Match 100.0%; Score 41; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASTTTNYT 8
 DB 1 ASTTTNYT 8
 RESULT 7
 AAR44336
 ID AAR44336 standard; peptide; 8 AA.
 XX
 AC AAR44336;
 XX
 DT 25-MAR-2003 (revised)
 DT 22-JUN-1994 (first entry)
 XX
 DE Peptide T related peptide for treating inflammation.
 XX
 KW Inflammation; antiinflammatory; immunomodulator; HTLV-1;
 KW multiple sclerosis; myopathy; chronic fatigue syndrome; toxic shock;
 KW arthritis; inflammatory bowel disease; host-versus-graft;
 KW graft-versus-host; transplant.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1..8
 FT /note= "one or more of these sites can be glycosylated"
 FT Misc-difference 1
 FT /note= "D-Ala"
 FT /note= "Thr-NH2"
 XX
 PN WO9320102-A1.
 XX
 PD 14-OCT-1993.
 XX
 PF 29-MAR-1993; 93WO-GB000649.
 XX
 PR 27-MAR-1992; 92US-00858832.
 PR 14-MAY-1992; 92DK-00000645.
 PR 17-JUL-1992; 92US-00915118.
 PR 09-DEC-1992; 92US-00987674.
 XX

PA (PEPT-) PEPTIDE TECHNOLOGY LTD.
 PA (DRUG-) DRUG ROYALTY CORP.
 XX
 PI Andersen AJ, Aston R, Carlen PL, Doob PR, Macfadden DK;
 PI Phipps DJ, Rathjen D, Widmer F;
 XX
 DR WPI; 1993-336832/42.
 XX
 PT New linear and cyclic peptide T analogues - used for treatment and
 FT prevention of inflammation, e.g. multiple sclerosis and myelopathies.
 XX
 PS Claim 7; Page 83; 103pp; English.
 XX
 CC The invention relates to a group of linear or cyclic tetra- to deca-
 CC peptides which include Peptide T and its analogues, where at least one of
 CC the amino acids may optionally have a monomeric or polymeric carbohydrate
 CC substituted onto a hydroxyl, amino and/or amido group. In one aspect the
 CC invention relates to the use of these peptides for the treatment of
 CC inflammation; and in another aspect it relates to those peptides in which
 CC the carbohydrate moiety is present (other than glycosylated Prototype
 CC Peptide T) as new chemical entities. The present peptide is one of 10
 CC specific peptides listed in a sub-claim for use in the treatment of
 CC inflammation. The peptides act as immunomodulators in the treatment and
 CC prevention of inflammation. In particular they can be used in the
 CC treatment of multiple sclerosis, myopathies (including HTLV-1 associated
 CC myopathies) and symptoms and diseases associated with chronic immune
 CC activation including chronic fatigue syndrome, toxic shock, arthritis,
 CC inflammatory bowel disease, and host-versus-graft and graft-versus-host
 CC responses in transplant recipients. (Updated on 25-MAR-2003 to correct PN
 CC field.)
 XX
 SQ Sequence 8 AA;
 Query Match 100.0%; Score 41; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASTTTNYT 8
 DB 1 ASTTTNYT 8
 RESULT 8
 AAR44334
 ID AAR44334 standard; peptide; 8 AA.
 XX
 AC AAR44334;
 XX
 DT 25-MAR-2003 (revised)
 DT 22-JUN-1994 (first entry)
 XX
 DE Peptide T related peptide for treating inflammation.
 XX
 KW Inflammation; antiinflammatory; immunomodulator; HTLV-1;
 KW multiple sclerosis; myopathy; chronic fatigue syndrome; toxic shock;
 KW arthritis; inflammatory bowel disease; host-versus-graft;
 KW graft-versus-host; transplant.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1..8
 FT /note= "one or more residues may be glycosylated"
 FT Misc-difference 1
 FT /note= "D-Ala"
 FT /note= "Thr-NH2"
 XX
 PN WO9320102-A1.
 XX
 PD 14-OCT-1993.
 XX
 PF 29-MAR-1993; 93WO-GB000649.

XX 27-MAR-1992; 92US-00858832.
 PR 14-MAY-1992; 92DK-00000645.
 PR 17-JUL-1992; 92US-00915118.
 PR 09-DEC-1992; 92US-00987674.
 XX (PEPT-) PEPTIDE TECHNOLOGY LTD.
 PA (DRUG-) DRUG ROYALTY CORP.
 XX Andersen AJ, Aston R, Carlen PL, Doob PR, Macfadden DK;
 PI Phipps DJ, Rathjen D, Widmer F;
 XX WPI; 1993-336832/42.
 XX New linear and cyclic peptide T analogues - used for treatment and
 PT prevention of inflammation, e.g. multiple sclerosis and myelopathies.
 XX Claim 7; Page 83; 103pp; English.
 XX The invention relates to a group of linear or cyclic tetra- to deca-
 CC peptides which include Peptide T and its analogues, where at least one of
 CC the amino acids may optionally have a monomeric or polymeric carbohydrate
 CC substituted onto a hydroxyl, amino and/or amide group. In one aspect the
 CC invention relates to the use of these peptides for the treatment of
 CC inflammation; and in another aspect it relates to those peptides in which
 CC the carbohydrate moiety is present (other than glycosylated prototype
 CC peptide T) as new chemical entities. The present peptide (Prototype
 CC peptide T) is one of 10 specific peptides listed in a sub-claim for use
 CC in the treatment of inflammation. The peptides act as immunomodulators in
 CC the treatment and prevention of inflammation. In particular they can be
 CC used in the treatment of multiple sclerosis, myopathies (including HTLV-1
 CC associated myopathies) and symptoms and diseases associated with chronic
 CC immune activation including chronic fatigue syndrome, toxic shock,
 CC arthritis, inflammatory bowel disease, and host-versus-graft and graft-
 CC versus-host responses in transplant recipients. (Updated on 25-MAR-2003
 CC to correct PN field.)
 XX Sequence 8 AA;
 SQ Query Match 100.0%; Score 41; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ASTTTNYT 8
 Db 1 ASTTTNYT 8
 RESULT 9
 AAR44335
 ID AAR44335 standard; peptide; 8 AA.
 AC AAR44335;
 XX 25-MAR-2003 (revised)
 DT 22-JUN-1994 (first entry)
 DE Peptide T related peptide for treating inflammation.
 XX Inflammation; antiinflammatory; immunomodulator; HTLV-1;
 KW multiple sclerosis; myopathy; chronic fatigue syndrome; toxic shock;
 KW arthritis; inflammatory bowel disease; host-versus-graft;
 KW graft-versus-host; transplant.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 1. .8
 FT /note= "one or more residues may be glycosylated"
 FN W09320102-A1.
 XX 14-OCT-1993.

XX 29-MAR-1993; 93WO-GB000649.
 XX 27-MAR-1992; 92US-00858832.
 PR 14-MAY-1992; 92DK-00000645.
 PR 17-JUL-1992; 92US-00915118.
 XX 09-DEC-1992; 92US-00987674.
 XX (PEPT-) PEPTIDE TECHNOLOGY LTD.
 PA (DRUG-) DRUG ROYALTY CORP.
 XX Andersen AJ, Aston R, Carlen PL, Doob PR, Macfadden DK;
 PI Phipps DJ, Rathjen D, Widmer F;
 XX WPI; 1993-336832/42.
 XX New linear and cyclic peptide T analogues - used for treatment and
 PT prevention of inflammation, e.g. multiple sclerosis and myelopathies.
 XX Claim 7; Page 83; 103pp; English.
 XX The invention relates to a group of linear or cyclic tetra- to deca-
 CC peptides which include Peptide T and its analogues, where at least one of
 CC the amino acids may optionally have a monomeric or polymeric carbohydrate
 CC substituted onto a hydroxyl, amino and/or amide group. In one aspect the
 CC invention relates to the use of these peptides for the treatment of
 CC inflammation; and in another aspect it relates to those peptides in which
 CC the carbohydrate moiety is present (other than glycosylated prototype
 CC peptide T) as new chemical entities. The present peptide is one of 10
 CC specific peptides listed in a sub-claim for use in the treatment of
 CC inflammation. The peptides act as immunomodulators in the treatment and
 CC prevention of inflammation. In particular they can be used in the
 CC treatment of multiple sclerosis, myopathies (including HTLV-1 associated
 CC myopathies) and symptoms and diseases associated with chronic immune
 CC activation including chronic fatigue syndrome, toxic shock, arthritis,
 CC inflammatory bowel disease, and host-versus-graft and graft-versus-host
 CC responses in transplant recipients. (Updated on 25-MAR-2003 to correct PN
 CC field.)
 XX Sequence 8 AA;
 SQ Query Match 100.0%; Score 41; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ASTTTNYT 8
 Db 1 ASTTTNYT 8
 RESULT 10
 AAR44342
 ID AAR44342 standard; peptide; 8 AA.
 AC AAR44342;
 XX 25-MAR-2003 (revised)
 DT 22-JUN-1994 (first entry)
 DE Peptide T related analogue 504 for treating inflammation.
 XX Inflammation; antiinflammatory; immunomodulator; HTLV-1;
 KW multiple sclerosis; myopathy; chronic fatigue syndrome; toxic shock;
 KW arthritis; inflammatory bowel disease; host-versus-graft;
 KW graft-versus-host; transplant.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 1. .8
 FT /note= "one or more sites may be glycosylated"
 FN Misc-difference 1
 FT /note= "D-Ala"

FT Misc-difference 3 /note= "D-Thr"
FT Misc-difference 5 /note= "D-Thr"
FT Misc-difference 8 /note= "D-Thr-NH2"
FT W09320102-A1.
XX 14-OCT-1993.
XX 29-MAR-1993; 93WO-GB000649.
XX 27-MAR-1992; 92US-00858832.
XX 14-MAY-1992; 92DK-00000645.
XX 17-JUL-1992; 92US-00915118.
XX 09-DEC-1992; 92US-00987674.
XX (PEPT-) PEPTIDE TECHNOLOGY LTD.
XX (DRUG-) DRUG ROYALTY CORP.
XX Andersen AJ, Aston R, Carlen PL, Doob PR, Macfadden DK;
XX Phipps DJ, Rathjen D, Widmer F;
XX WPI; 1993-336832/42.
XX New linear and cyclic peptide T analogues - used for treatment and
XX prevention of inflammation, e.g. multiple sclerosis and myelopathies.
XX Claim 7; Page 84; 103pp; English.
XX The invention relates to a group of linear or cyclic tetra- to deca-
XX peptides which include peptide T and its analogues, where at least one of
XX the amino acids may optionally have a monomeric or polymeric carbohydrate
XX substituted onto a hydroxyl, amino and/or amido group. In one aspect the
XX invention relates to the use of these peptides for the treatment of
XX inflammation; and in another aspect it relates to those peptides in which
XX the carbohydrate moiety is present (other than glycosylated Prototypic
XX Peptide T) as new chemical entities. The present peptide is one of 10
XX specific peptides listed in a sub-claim for use in the treatment of
XX inflammation. The peptides act as immunomodulators in the treatment and
XX prevention of inflammation. In particular they can be used in the
XX treatment of multiple sclerosis, myopathies (including HTLV-1 associated
XX myopathies) and symptoms and diseases associated with chronic immune
XX activation including chronic fatigue syndrome, toxic shock, arthritis,
XX inflammatory bowel disease, and host-versus-graft and graft-versus-host
XX responses in transplant recipients. (Updated on 25-MAR-2003 to correct PN
XX field.)
XX Sequence 8 AA;
Query Match 100.0%; Score 41; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ASTTTNYT 8
Db 1 ASTTTNYT 8
RESULT 11
AAR47827
ID AAR47827 standard; peptide; 8 AA.
XX AAR47827;
XX 25-MAR-2003 (revised)
XX 22-JUN-1994 (first entry)
XX Peptide T for intranasal therapy of immunodeficiency disease.
XX TSP; tropical spastic paresis; HTLV-1, neuromotor; cognitive; HIV; gp120;
XX replication; AIDS; vaccine.
XX

XX OS Synthetic.
XX EP579363-A1.
XX 19-JAN-1994.
XX 20-MAY-1993; 93EP-00303919.
XX 15-JUN-1992; 92US-00898691.
XX (ADPE-) ADVANCED PEPTIDES & BIOTECHNOLOGY SCI.
XX Bert CB, Ruff MR;
XX WPI; 1994-017880/03.
XX Compens. to treat tropical spastic paresis and immunodeficiency - contain
XX peptide to block receptor binding or viruses.
XX Claim 6; Page 16; 22pp; English.
XX The peptide, a peptide T analogue, has a high threonine content. The
XX peptide binds to T4 receptors and is useful for intranasal admin. to
XX prevent viral infectivity in mammals by viruses which bind to the T4
XX receptors. The peptide is believed to act as a competitive blocking
XX agent. It inhibits binding of HIV gp 120 to brain tissue and inhibits HIV
XX replication in cell culture. It can be used to alleviate the symptoms of
XX AIDS. It can also be used to treat TSP, resulting in loss of paralysis
XX and improvement in bladder and sexual function and decreased
XX deterioration of memory and attention. The peptide may also be used in
XX vaccines to prevent the transmission of HIV. See also AAR47826-39.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 8 AA;
Query Match 100.0%; Score 41; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ASTTTNYT 8
Db 1 ASTTTNYT 8
RESULT 12
AAR59486
ID AAR59486 standard; peptide; 8 AA.
XX AAR59486;
XX 25-MAR-2003 (revised)
XX 26-NOV-1995 (first entry)
XX Prototypic Peptide T used for treating eczema.
XX peptide T; prototypic; octapeptide; eczema; dermatitis.
XX Synthetic.
XX Key Location/Qualifiers
XX Misc-difference 1 /note= "D-form residue"
XX Modified-site 8 /note= "Thr-NH2"
XX WO9508338-A1.
XX 30-MAR-1995.
XX 26-SEP-1994; 94WO-AU000574.
XX 24-SEP-1993; 93AU-00001449.
XX

```
XX (PEPT-) PEPTIDE TECHNOLOGY LTD.
XX Michaelis J, Trigg TE;
XX WPI; 1995-139385/18.
XX Treatment or prevention of eczema or dermatitis, esp. atopic eczema -
XX using specified peptide, e.g the octa:peptide Peptide T.
XX Claim 3; Page 14; 20pp; English.
XX The invention relates to a new method of treating eczema or dermatitis
XX using Peptide T and its variants. The present sequence is that of
XX prototype peptide T which is one of the preferred peptides. (Updated
XX on 25-MAR-2003 to correct PN field.)
XX Sequence 8 AA;
XX Query Match 100.0%; Score 41; DB 2; Length 8;
XX Best Local Similarity 100.0%; Pred. No. 1.8e+06;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 ASTTTNYT 8
XX DB 1 ASTTTNYT 8
XX RESULT 14
XX AAR59493
XX ID AAR59493 standard; peptide; 8 AA.
XX AC AAR59493;
XX DT 25-MAR-2003 (revised)
XX DT 26-NOV-1995 (first entry)
XX DE Peptide T variant used for treating eczema.
XX KW peptide T; prototype; octapeptide; eczema; dermatitis.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 1 /note= "D-form residue"
XX FT Misc-difference 3 /note= "D-form residue"
XX FT Misc-difference 5 /note= "D-form residue"
XX FT Misc-difference 8 /note= "D-form residue"
XX FT Misc-difference 8 /note= "D-Thr-NH2"
XX PN WO9508338-A1.
XX DT 30-MAR-1995.
XX DT 26-SEP-1994; 94WO-AU000574.
XX PR 24-SEP-1993; 93AU-00001449.
XX PA (PEPT-) PEPTIDE TECHNOLOGY LTD.
XX PI Michaelis J, Trigg TE;
XX DR WPI; 1995-139385/18.
XX XX Treatment or prevention of eczema or dermatitis, esp. atopic eczema -
XX using specified peptide, e.g the octa:peptide Peptide T.
XX Claim 3; Page 15; 20pp; English.
XX The invention relates to a new method of treating eczema or dermatitis
XX using Peptide T and its variants. The present sequence represents one of
XX the preferred variants. (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 8 AA;
XX Query Match 100.0%; Score 41; DB 2; Length 8;
XX Best Local Similarity 100.0%; Pred. No. 1.8e+06;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 ASTTTNYT 8
XX DB 1 ASTTTNYT 8
XX RESULT 15
XX AAR61488
XX ID AAR61488 standard; peptide; 8 AA.
XX AC AAR61488;
XX XX
```

```
XX (PEPT-) PEPTIDE TECHNOLOGY LTD.
XX Michaelis J, Trigg TE;
XX WPI; 1995-139385/18.
XX Treatment or prevention of eczema or dermatitis, esp. atopic eczema -
XX using specified peptide, e.g the octa:peptide Peptide T.
XX Claim 3; Page 14; 20pp; English.
XX The invention relates to a new method of treating eczema or dermatitis
XX using Peptide T and its variants. The present sequence is that of
XX prototype peptide T which is one of the preferred peptides. (Updated
XX on 25-MAR-2003 to correct PN field.)
XX Sequence 8 AA;
XX Query Match 100.0%; Score 41; DB 2; Length 8;
XX Best Local Similarity 100.0%; Pred. No. 1.8e+06;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 ASTTTNYT 8
XX DB 1 ASTTTNYT 8
XX RESULT 13
XX AAR59487
XX ID AAR59487 standard; peptide; 8 AA.
XX AC AAR59487;
XX DT 25-MAR-2003 (revised)
XX DT 26-NOV-1995 (first entry)
XX DE Peptide T variant used for treating eczema.
XX KW peptide T; prototype; octapeptide; eczema; dermatitis.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 1 /note= "L- or D-form residue"
XX PN WO9508338-A1.
XX DT 30-MAR-1995.
XX DT 26-SEP-1994; 94WO-AU000574.
XX PR 24-SEP-1993; 93AU-00001449.
XX PA (PEPT-) PEPTIDE TECHNOLOGY LTD.
XX PI Michaelis J, Trigg TE;
XX DR WPI; 1995-139385/18.
XX XX Treatment or prevention of eczema or dermatitis, esp. atopic eczema -
XX using specified peptide, e.g the octa:peptide Peptide T.
XX Claim 3; Page 14; 20pp; English.
XX The invention relates to a new method of treating eczema or dermatitis
XX using Peptide T and its variants. The present sequence represents one of
XX the preferred variants. (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 8 AA;
XX Query Match 100.0%; Score 41; DB 2; Length 8;
```

Mon Mar 21 08:14:28 2005

DT 25-MAR-2003 (revised)
 XX 26-NOV-1995 (first entry)
 DE Peptide T variant for use in treating Crohn's disease.
 XX peptide T; prototype; octapeptide; Crohn's disease; ulcerative; colitis.
 KW peptide T; prototype; octapeptide; Crohn's disease; ulcerative; colitis.
 XX Synthetic.
 OS
 XX
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "D-form residue"
 FT Misc-difference 3 /note= "D-form residue"
 FT Misc-difference 5 /note= "D-form residue"
 FT Misc-difference 8 /note= "D-form residue"
 FT Misc-difference 8 /note= "D-Thr-NH2"
 FT
 XX WO9508339-A1.
 PN
 XX
 XX
 PD 30-MAR-1995.
 XX
 PF 26-SEP-1994; 94WO-AU000575.
 XX
 PR 24-SEP-1993; 93AU-00001448.
 XX
 PA (PEPT-) PEPTIDE TECHNOLOGY LTD.
 XX
 XX Michaelis J, Sleight MJ;
 PI
 XX WPI; 1995-139386/18.
 DR
 XX Treatment or prevention of Crohn's disease or ulcerative colitis - by use
 PT of specified peptide, e.g. Peptide T.
 PT
 PS Claim 3; Page 15; 24pp; English.
 XX
 CC The invention relates to a new method of treating or preventing Crohn's
 CC disease and/or ulcerative colitis using Peptide T and its variants. The
 CC present sequence is one of the preferred peptides. (Updated on 25-MAR-
 CC 2003 to correct PN field.)
 XX
 XX Sequence 8 AA;
 SQ
 Query Match 100.0%; Score 41; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ASTTTNYT 8
 DB 1 ASTTTNYT 8
 Search completed: March 18, 2005, 19:10:10
 Job time : 80.16 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 18:59:50 ; Search time 19.52 Seconds
(without alignments)
30.594 Million cell updates/sec

Title: US-10-773-274A-1

Perfect score: 41

Sequence: 1 ASTTNTYT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41	100.0	8	1	US-07-920-597-4
2	41	100.0	8	1	US-08-302-829-3
3	41	100.0	8	1	US-08-302-829-5
4	41	100.0	8	1	US-08-302-829-10
5	41	100.0	8	1	US-08-488-252-32
6	41	100.0	8	1	US-08-481-840A-1
7	41	100.0	8	1	US-08-481-840A-2
8	41	100.0	8	1	US-08-481-840A-3
9	41	100.0	8	1	US-08-403-718-1
10	41	100.0	8	1	US-08-403-718-6
11	41	100.0	8	1	US-08-619-462-1
12	41	100.0	8	1	US-08-619-462-6
13	41	100.0	8	2	US-08-747-137-20
14	41	100.0	8	3	US-08-082-837A-3
15	41	100.0	8	3	US-08-082-837A-5
16	41	100.0	8	3	US-09-082-837A-10
17	41	100.0	8	3	US-09-421-845-3
18	41	100.0	8	3	US-09-421-845-5
19	41	100.0	8	3	US-09-421-845-10
20	41	100.0	9	1	US-08-201-046A-29
21	41	100.0	509	3	US-08-472-240A-20
22	41	100.0	610	3	US-09-257-490-12
23	41	100.0	855	3	US-07-956-483-15
24	41	100.0	887	3	US-08-472-240A-6
25	41	100.0	1018	5	PCT-US93-11703-75
26	37	90.2	8	6	5276016-2
27	37	90.2	8	6	5276016-2

28	33	80.5	116	4	US-09-248-796A-21372	Sequence 21372, A
29	33	80.5	468	4	US-09-270-767-44250	Sequence 44250, A
30	32	78.0	6	1	US-08-403-718-5	Sequence 5, Appli
31	32	78.0	6	1	US-08-619-462-5	Sequence 5, Appli
32	32	78.0	6	6	5276016-1	Patent No. 5276016
33	32	78.0	6	6	5276016-1	Patent No. 5276016
34	32	78.0	8	1	US-08-554-758-1	Sequence 1, Appli
35	32	78.0	8	1	US-08-302-829-2	Sequence 2, Appli
36	32	78.0	8	1	US-08-302-829-11	Sequence 11, Appl
37	32	78.0	8	1	US-08-403-718-4	Sequence 4, Appli
38	32	78.0	8	1	US-08-619-462-4	Sequence 4, Appli
39	32	78.0	8	3	US-09-082-837A-2	Sequence 2, Appli
40	32	78.0	8	3	US-09-082-837A-11	Sequence 11, Appl
41	32	78.0	8	3	US-09-421-845-2	Sequence 2, Appli
42	32	78.0	8	3	US-09-421-845-11	Sequence 11, Appl
43	32	78.0	10	1	US-08-385-443-1	Sequence 1, Appli
44	32	78.0	135	4	US-09-248-796A-22190	Sequence 22190, A
45	32	78.0	194	4	US-09-248-796A-25232	Sequence 25232, A

ALIGNMENTS

RESULT 1
US-07-920-597-4
; Sequence 4, Application US/07920597
; Patent No. 5447915
; GENERAL INFORMATION:
; APPLICANT: Schreiber, Stuart
; APPLICANT: Burakoff, Steven
; TITLE OF INVENTION: Terminally-Blocked Antiviral Peptides
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Aligretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/920,597
; FILING DATE: 19920828
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO pct/us91/01142
; FILING DATE: 28-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5447915nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 91,174-E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-920-597-4

Query Match 100.0%; Score 41; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ASTTNTYT 8

Db 1 ASTTTNYT 8
 |||||
 RESULT 2
 US-08-302-829-3
 ; Sequence 3, Application US/08302829
 ; Patent No. 5756449
 ; GENERAL INFORMATION:
 ; APPLICANT: ANDERSEN, Anders J
 ; APPLICANT: ASTON, Roger
 ; APPLICANT: CARLEN, Peter L
 ; APPLICANT: DOOB, Penelope R
 ; APPLICANT: MacFADDAEN, Douglas K
 ; APPLICANT: PHIPPS, David J
 ; APPLICANT: RATHJEN, Deborah
 ; APPLICANT: WIDMER, Fred
 ; TITLE OF INVENTION: Peptide T and Related Peptides in the
 ; TITLE OF INVENTION: Treatment of Inflammation, Including Multiple Sclerosis
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BANNER & ALLEGRETTI, LTD.
 ; STREET: 10 S. Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/302.829
 ; FILING DATE: 29-MAR-1993
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/GB93/00649
 ; FILING DATE: 29-MAR-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/987,674
 ; FILING DATE: 09-DEC-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/915,118
 ; FILING DATE: 17-JUL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: DK 645/92
 ; FILING DATE: 14-MAY-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/858,832
 ; FILING DATE: 27-MAR-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26949
 ; REFERENCE/DOCKET NUMBER: 94,772
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 8 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: both
 ; MOLECULE TYPE: peptide
 ; FEATURE:
 ; NAME/KEY: Peptide
 ; LOCATION: 1..8
 ; OTHER INFORMATION: /note="Ala 1 is D-Ala"
 US-08-302-829-3
 Query Match 100.0%; Score 41; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 4.1e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ASTTTNYT 8
 |||||
 Db 1 ASTTTNYT 8
 |||||
 RESULT 3
 US-08-302-829-5
 ; Sequence 5, Application US/08302829
 ; Patent No. 5756449
 ; GENERAL INFORMATION:
 ; APPLICANT: ANDERSEN, Anders J
 ; APPLICANT: ASTON, Roger
 ; APPLICANT: CARLEN, Peter L
 ; APPLICANT: DOOB, Penelope R
 ; APPLICANT: MacFADDAEN, Douglas K
 ; APPLICANT: PHIPPS, David J
 ; APPLICANT: RATHJEN, Deborah
 ; APPLICANT: WIDMER, Fred
 ; TITLE OF INVENTION: Peptide T and Related Peptides in the
 ; TITLE OF INVENTION: Treatment of Inflammation, Including Multiple Sclerosis
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BANNER & ALLEGRETTI, LTD.
 ; STREET: 10 S. Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/302.829
 ; FILING DATE: 29-MAR-1993
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/GB93/00649
 ; FILING DATE: 29-MAR-1993
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/987,674
 ; FILING DATE: 09-DEC-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/915,118
 ; FILING DATE: 17-JUL-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: DK 645/92
 ; FILING DATE: 14-MAY-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/858,832
 ; FILING DATE: 27-MAR-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McDonnell, John J
 ; REGISTRATION NUMBER: 26949
 ; REFERENCE/DOCKET NUMBER: 94,772
 ; TELEPHONE: 312-715-1000
 ; TELEFAX: 312-715-1234
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 8 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: not relevant
 ; TOPOLOGY: circular
 ; MOLECULE TYPE: peptide
 US-08-302-829-5
 Query Match 100.0%; Score 41; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 4.1e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASTTTNYT 8
| | | | |
Db 1 ASTTTNYT 8

RESULT 4

US-08-302-829-10
; Sequence 10, Application US/08302829
; Patent No. 5756449
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Anders J
; APPLICANT: ASTON, Roger
; APPLICANT: CARLEN, Peter L
; APPLICANT: DOOB, Penelope R
; APPLICANT: MacFADDEN, Douglas K
; APPLICANT: PHIPPS, David J
; APPLICANT: RATHJEN, Deborah
; APPLICANT: WIDMER, Fred
; TITLE OF INVENTION: Peptide T and Related Peptides in the
; TREATMENT OF Inflammation, Including Multiple Sclerosis
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER & ALLEGRETTI, LTD.
; STREET: 10 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/302,829
; APPLICATION NUMBER: US/08/302,829
; FILING DATE: 29-MAR-1993
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB93/00649
; FILING DATE: 29-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/987,674
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/915,118
; FILING DATE: 17-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 645/92
; FILING DATE: 14-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/858,832
; FILING DATE: 27-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: McDowell, John J
; REGISTRATION NUMBER: 26949
; REFERENCE/DOCKET NUMBER: 94,772
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..8
; OTHER INFORMATION: /note= "aa1 is D-Ala, aa3 is D-Thr,
; OTHER INFORMATION: aa5 is D-Thr, aa8 is D-Thr"
US-08-302-829-10

Query Match 100.0%; Score 41; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASTTTNYT 8
| | | | |
Db 1 ASTTTNYT 8

RESULT 5

US-08-488-252-32
; Sequence 32, Application US/08488252
; Patent No. 5763160
; GENERAL INFORMATION:
; APPLICANT: Chang Yi Wang
; TITLE OF INVENTION: SYNTHETIC PEPTIDES AND PROCESS
; OF USING SAME FOR THE DETECTION OF ANTIBODIES TO
; HUMAN IMMUNODEFICIENCY VIRUS (HIV) GP120 ENVELOPE
; TITLE OF INVENTION: PROTEIN, DIAGNOSIS OF AIDS AND PRE-AIDS CONDITIONS
; TITLE OF INVENTION: AND AS VACCINES
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVE.
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA: US/08/488,252
; APPLICATION NUMBER: US/08/488,252
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08\326,676
; FILING DATE: 07-Jun-1995
; APPLICATION NUMBER: 07\726,605
; FILING DATE: 09-July-1991
; APPLICATION NUMBER: 07\663,262
; FILING DATE: 01-Mar-1991
; APPLICATION NUMBER: 07\155,321
; FILING DATE: 12-Feb-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C. H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4004 USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: Amino acids
; STRANDEDNESS:
; TOPOLOGY: Unknown
US-08-488-252-32

Query Match 100.0%; Score 41; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASTTTNYT 8
| | | | |
Db 1 ASTTTNYT 8

RESULT 6

```
US-08-481-840A-1
; Sequence 1, Application US/08481840A
; Patent No. 5763406
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, Ove
; TITLE OF INVENTION: Peptides for Treatment
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, Ltd.
; STREET: 10 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk, 3.50 inch
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481.840A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/107,777
; FILING DATE: 24-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Iwanicki, John P
; REGISTRATION NUMBER: 34,628
; REFERENCE/DOCKET NUMBER: 93,741-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-345-9100
; TELEFAX: 617-345-9111
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: Linear
; MOLECULE TYPE: peptide
US-08-481-840A-1
Query Match 100.0%; Score 41; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTTTNYT 8
Db 1 ASTTTNYT 8

RESULT 7
US-08-481-840A-2
; Sequence 2, Application US/08481840A
; Patent No. 5763406
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, Ove
; TITLE OF INVENTION: Peptides for Treatment
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, Ltd.
; STREET: 10 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk, 3.50 inch
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481.840A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/107,777
; FILING DATE: 24-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Iwanicki, John P
; REGISTRATION NUMBER: 34,628
; REFERENCE/DOCKET NUMBER: 93,741-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-345-9100
; TELEFAX: 617-345-9111
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: Linear
; MOLECULE TYPE: peptide
US-08-481-840A-1
Query Match 100.0%; Score 41; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTTTNYT 8
Db 1 ASTTTNYT 8

RESULT 8
US-08-481-840A-3
; Sequence 3, Application US/08481840A
; Patent No. 5763406
; GENERAL INFORMATION:
; APPLICANT: PEDERSEN, Ove
; TITLE OF INVENTION: Peptides for Treatment
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, Ltd.
; STREET: 10 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk, 3.50 inch
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481.840A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/107,777
; FILING DATE: 24-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Iwanicki, John P
; REGISTRATION NUMBER: 34,628
; REFERENCE/DOCKET NUMBER: 93,741-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-345-9100
; TELEFAX: 617-345-9111
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: Linear
```


TELEFAX: (704) 816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-462-1

Query Match 100.0%; Score 41; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASTTNYT 8
| | | | |
Db 1 ASTTNYT 8

RESULT 12
US-08-619-462-6
; Sequence 6, Application US/08619462
; Patent No. 5798335
; GENERAL INFORMATION:
; APPLICANT: MICHAELIS, JURGEN
; APPLICANT: Trigg, Timothy Elliot
; TITLE OF INVENTION: METHOD FOR THE TREATMENT OR PREVENTION OF
; TITLE OF INVENTION: ECZEMA/DERMATITIS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,462
; FILING DATE: 07-MAY-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MITCHARD, LEONARD C.
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 47-84
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (704) 816-4100
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-619-462-6

Query Match 100.0%; Score 41; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASTTNYT 8
| | | | |
Db 1 ASTTNYT 8

RESULT 13
US-08-747-137-20
; Sequence 20, Application US/08747137

Patent No. 5945033
; GENERAL INFORMATION:
; APPLICANT: YEN, Richard C.K.
; TITLE OF INVENTION: NON-CROSSLINKED PROTEIN PARTICLES FOR
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC USE
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/747,137
; FILING DATE: 12-NOV-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,546
; FILING DATE: 14-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/069,831
; FILING DATE: 01-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/959,560
; FILING DATE: 13-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/641,720
; FILING DATE: 15-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 016197-000840US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
US-08-747-137-20

Query Match 100.0%; Score 41; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASTTNYT 8
| | | | |
Db 1 ASTTNYT 8

RESULT 14
US-09-082-837A-3
; Sequence 3, Application US/09082837A
; Patent No. 6011014
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Anders J
; APPLICANT: ASTON, Roger
; APPLICANT: CARLEN, Peter L
; APPLICANT: DOOB, Penelope R
; APPLICANT: MacFADDAEN, Douglas K
; APPLICANT: PHIPPS, David J
; APPLICANT: RATHJEN, Deborah
; APPLICANT: WIDMER, Fred
; TITLE OF INVENTION: Peptide T and Related Peptides in the Treatment of
; TITLE OF INVENTION: Inflammation, Including Multiple sclerosis
; NUMBER OF SEQUENCES: 11

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BANNER & WITCOFF, LTD.
;; STREET: 10 S. Wacker Drive
;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: USA
;; ZIP: 60606
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/082,837A
;; FILING DATE: 21-JUN-1998
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/302,829
;; FILING DATE: 29-MAR-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/GB93/00649
;; FILING DATE: 29-MAR-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/987,674
;; FILING DATE: 09-DEC-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/915,118
;; FILING DATE: 17-JUL-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: DK 645/92
;; FILING DATE: 14-MAY-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/858,832
;; FILING DATE: 27-MAR-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Drehkoff, W. Dennis
;; REGISTRATION NUMBER: 27193
;; REFERENCE/DOCKET NUMBER: 94, 772-A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312-715-1000
;; TELEFAX: 312-715-1234
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 8 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: both
;; MOLECULE TYPE: peptide
;; FEATURE:
;; NAME/KEY: Peptide
;; LOCATION: 1..8
;; OTHER INFORMATION: /note= "Ala 1 is D-Ala"
US-09-082-837A-3

Query Match 100.0%; Score 41; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASTTTNYT 8
Db 1 ASTTTNYT 8

RESULT 15

US-09-082-837A-5
; Sequence 5, Application US/09082837A
; Patent No. 6011014
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Anders J
; APPLICANT: ASTON, Roger
; APPLICANT: CARLEN, Peter L
; APPLICANT: DOOB, Penelope R
; APPLICANT: MacFADDAEN, Douglas K

;; APPLICANT: PHIPPS, David J
;; APPLICANT: RATHJEN, Deborah
;; APPLICANT: WIDMER, Fred
;; TITLE OF INVENTION: Peptide T and Related Peptides in the Treatment of
;; TITLE OF INVENTION: Inflammation, Including Multiple sclerosis
;; NUMBER OF SEQUENCES: 11
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BANNER & WITCOFF, LTD.
;; STREET: 10 S. Wacker Drive
;; CITY: Chicago
;; STATE: Illinois
;; COUNTRY: USA
;; ZIP: 60606
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/082,837A
;; FILING DATE: 21-JUN-1998
;; CLASSIFICATION: 514
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/302,829
;; FILING DATE: 29-MAR-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/GB93/00649
;; FILING DATE: 29-MAR-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/987,674
;; FILING DATE: 09-DEC-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/915,118
;; FILING DATE: 17-JUL-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: DK 645/92
;; FILING DATE: 14-MAY-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/858,832
;; FILING DATE: 27-MAR-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Drehkoff, W. Dennis
;; REGISTRATION NUMBER: 27193
;; REFERENCE/DOCKET NUMBER: 94, 772-A
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312-715-1000
;; TELEFAX: 312-715-1234
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 8 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: not relevant
;; TOPOLOGY: circular
;; MOLECULE TYPE: peptide
US-09-082-837A-5

Query Match 100.0%; Score 41; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASTTTNYT 8
Db 1 ASTTTNYT 8

Search completed: March 18, 2005, 19:12:09
Job time : 20:52 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 19:02:16 ; Search time 58.88 Seconds
(without alignments)
44.910 Million cell updates/sec

Title: US-10-773-274A-1

Perfect score: 41

Sequence: 1 ASSTTNYT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1401741 seqs, 330541175 residues

Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pdb.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pdb.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pdb.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pdb.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pdb.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pdb.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pdb.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pdb.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pdb.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pdb.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pdb.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pdb.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pdb.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pdb.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pdb.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pdb.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pdb.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pdb.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pdb.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	9	14	US-10-062-710-230
2	41	100.0	451	14	US-10-361-849-15
3	41	100.0	509	16	US-10-872-198-124
4	35	85.4	121	15	US-10-424-599-166162
5	34	82.9	66	9	US-09-925-299-1516
6	34	82.9	66	10	US-09-925-299-1516
7	34	82.9	104	16	US-10-767-701-40122
8	34	82.9	358	15	US-10-449-857A-88
9	34	82.9	759	15	US-10-618-581-38
10	34	82.9	759	16	US-10-451-467A-364
11	33	80.5	118	15	US-10-424-599-207321
12	33	80.5	233	15	US-10-424-599-263645
13	33	80.5	237	15	US-10-425-114-40777

14	33	80.5	476	16	US-10-767-701-46985	Sequence 46985, A
15	33	80.5	483	15	US-10-425-114-43218	Sequence 43218, A
16	33	80.5	483	15	US-10-425-114-63234	Sequence 63234, A
17	32	78.0	40	9	US-09-764-847-951	Sequence 951, App
18	32	78.0	40	14	US-10-092-154-951	Sequence 951, App
19	32	78.0	92	15	US-10-424-599-223447	Sequence 223447, App
20	32	78.0	180	9	US-09-910-150-29	Sequence 29, Appl
21	32	78.0	180	15	US-10-377-097-29	Sequence 29, Appl
22	32	78.0	194	9	US-09-910-150-28	Sequence 28, Appl
23	32	78.0	194	15	US-10-377-097-28	Sequence 28, Appl
24	32	78.0	267	14	US-10-029-386-31950	Sequence 31950, A
25	32	78.0	269	16	US-10-437-963-176857	Sequence 176857, App
26	32	78.0	287	9	US-09-244-984-5	Sequence 5, Appli
27	32	78.0	287	14	US-10-057-321-5	Sequence 5, Appli
28	32	78.0	312	15	US-10-425-114-55428	Sequence 55428, A
29	32	78.0	318	17	US-10-774-355A-1590	Sequence 1590, Ap
30	32	78.0	343	15	US-10-424-599-234889	Sequence 234889, App
31	32	78.0	694	15	US-10-295-027-356	Sequence 356, App
32	32	78.0	698	9	US-09-727-169-4	Sequence 4, Appli
33	32	78.0	698	9	US-09-726-968-4	Sequence 4, Appli
34	32	78.0	698	14	US-10-281-478-2	Sequence 2, Appli
35	32	78.0	698	14	US-10-423-729-4	Sequence 4, Appli
36	32	78.0	702	16	US-10-437-963-123816	Sequence 123816, App
37	32	78.0	800	15	US-10-369-493-8983	Sequence 8983, Ap
38	32	78.0	824	9	US-09-727-169-2	Sequence 2, Appli
39	32	78.0	824	9	US-09-726-968-2	Sequence 2, Appli
40	32	78.0	824	14	US-10-281-478-1	Sequence 1, Appli
41	32	78.0	824	14	US-10-423-729-2	Sequence 2, Appli
42	32	78.0	824	15	US-10-295-027-354	Sequence 354, App
43	32	78.0	2386	14	US-10-156-761-7751	Sequence 7751, Ap
44	31	75.6	8	15	US-10-149-135-242	Sequence 242, App
45	31	75.6	8	15	US-10-149-135-621	Sequence 621, App

ALIGNMENTS

RESULT 1
US-10-062-710-230
; Sequence 230, Application US/10062710
; Publication No. US20030049253A1
; GENERAL INFORMATION:
; APPLICANT: Chu, Yong-Liang
; APPLICANT: Qiu, Jian-Tai
; APPLICANT: Li, Frank Q.
; TITLE OF INVENTION: Polymeric Conjugates for Delivery of
; TITLE OF INVENTION: MHC-Recognized Epitopes
; TITLE OF INVENTION: Via Peptide Vaccines
; FILE REFERENCE: 3781-001-27
; CURRENT APPLICATION NUMBER: US/10/062,710
; PRIOR FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: US 60/310,498
; PRIOR FILING DATE: 2001-08-08
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 230
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Epitopes of anti-HIV proteins
US-10-062-710-230

Query Match 100.0%; Score 41; DB 14; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.3e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASSTTNYT 8

Db 2 ASSTTNYT 9

RESULT 2

```
US-10-361-849-15
; Sequence 15, Application US/10361849
; Publication No. US20030170619A1
; GENERAL INFORMATION:
; APPLICANT: Hara, Toshio
; TITLE OF INVENTION: Nucleic Acid Capable of Promoting Gene Expression
; FILE REFERENCE: 027847, 00102US01
; CURRENT APPLICATION NUMBER: US/10/361,849
; CURRENT FILING DATE: 2003-02-11
; PRIOR FILING DATE: 2002-02-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus
US-10-361-849-15

Query Match      100.0%; Score 41; DB 14; Length 451;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ASTTTNYT 8
Db      127 ASTTTNYT 134

RESULT 3
US-10-872-198-124
; Sequence 124, Application US/10872198
; Publication No. US20050002897A1
; GENERAL INFORMATION:
; APPLICANT: Ulrich HAUPTS
; APPLICANT: Andre KOLTERMANN
; APPLICANT: Andreas SCHEIDIG
; APPLICANT: Christian VOETSMIER
; APPLICANT: Ulrich Ketting
; TITLE OF INVENTION: NEW BIOLOGICAL ENTITIES AND USE THEREOF
; FILE REFERENCE: 04156.0002U4
; CURRENT APPLICATION NUMBER: US/10/872,198
; CURRENT FILING DATE: 2004-06-18
; PRIOR FILING DATE: 2004-02-11
; PRIOR FILING DATE: 2003-11-25
; PRIOR FILING DATE: 2004-02-11
; PRIOR FILING DATE: 2003-11-11
; PRIOR FILING DATE: 2003-11-10
; PRIOR FILING DATE: 2003-06-18
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 124
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-872-198-124

Query Match      100.0%; Score 41; DB 16; Length 509;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ASTTTNYT 8
Db      185 ASTTTNYT 192

RESULT 4
US-10-424-599-166162
```

```
; Sequence 166162, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 166162
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_12105C.1.pep
US-10-424-599-166162

Query Match      85.4%; Score 35; DB 15; Length 121;
Best Local Similarity 87.5%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ASTTTNYT 8
Db      83 ASITTNYT 90

RESULT 5
US-09-925-299-1516
; Sequence 1516, Application US/09925299
; Patent No. US20020055827A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1516
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (33)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (47)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (52)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1516
```

Query Match 82.9%; Score 34; DB 9; Length 66;
Best Local Similarity 75.0%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ASTTTNT 8
Db 57 SKTTTNT 64

RESULT 6
US-09-925-299-1516
; Sequence 1516, Application US/09925299
; Publication No. US20030040617A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1516
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (33)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (47)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (52)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1516

Query Match 82.9%; Score 34; DB 10; Length 66;
Best Local Similarity 75.0%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ASTTTNT 8
Db 57 SKTTTNT 64

RESULT 7
US-10-767-701-40122
; Sequence 40122, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement

; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 40122
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(104)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C5227_1.pep
US-10-767-701-40122

Query Match 82.9%; Score 34; DB 16; Length 104;
Best Local Similarity 85.7%; Pred. No. 93;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 STTTNT 8
Db 17 NTTTNT 23

RESULT 8
US-10-449-857A-88
; Sequence 88, Application US/10449857A
; Publication No. US20040043931A1
; GENERAL INFORMATION:
; APPLICANT: Hersberg, Robert M.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Raodoh
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF INFLAMMATORY BOWEL DISEASE
; FILE REFERENCE: 594C1
; CURRENT APPLICATION NUMBER: US/10/449,857A
; CURRENT FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 88
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Unknown Bacterium
US-10-449-857A-88

Query Match 82.9%; Score 34; DB 15; Length 358;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ASTTTNT 8
Db 238 SDTTTNT 245

RESULT 9
US-10-618-581-38
; Sequence 38, Application US/10618581
; Publication No. US20040077524A1
; GENERAL INFORMATION:
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: METHOD FOR SELECTIVELY INHIBITING FUNGAL GROWTH
; FILE REFERENCE: 034536/0323
; CURRENT APPLICATION NUMBER: US/10/618,581
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,624
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent In Ver. 2.1

```
; SEQ ID NO 38
; LENGTH: 759
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-618-581-38

Query Match      82.9%; Score 34; DB 15; Length 759;
Best Local Similarity 85.7%; Pred. No. 6.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 STTTNYT 8
Db      64 TTTTNYT 70

RESULT 10
US-10-451-467A-364
; Sequence 364, Application US/10451467A
; Publication No. US20040161840A1
; GENERAL INFORMATION:
; APPLICANT: CONTRERAS, ROLAND HENRI
; APPLICANT: EBERHARDT, INES
; APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS
; APPLICANT: REEKMAN, RIEKA JOSEPHINA
; TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN
; TITLE OF INVENTION: YEAST AND FUNGI
; FILE REFERENCE: JAB-1667
; CURRENT APPLICATION NUMBER: US/10/451,467A
; CURRENT FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: EP 00870318.3
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: EP 01870002.1
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: EP 01870003.9
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 732
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 364
; LENGTH: 759
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-451-467A-364

Query Match      82.9%; Score 34; DB 16; Length 759;
Best Local Similarity 85.7%; Pred. No. 6.9e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 STTTNYT 8
Db      64 TTTTNYT 70

RESULT 11
US-10-424-599-207321
; Sequence 207321, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 207321
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_29236C.1.pap
US-10-424-599-207321

Query Match      80.5%; Score 33; DB 15; Length 118;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 STTTNYT 8
Db      108 STSTNYT 114

RESULT 12
US-10-424-599-263645
; Sequence 263645, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 263645
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_80091C.1.pap
US-10-424-599-263645

Query Match      80.5%; Score 33; DB 15; Length 233;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ASSTTNYT 8
Db      113 ASSTTNYT 120

RESULT 13
US-10-425-114-40777
; Sequence 40777, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 40777
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3028-028-G1_FLI.pap
US-10-425-114-40777

Query Match      80.5%; Score 33; DB 15; Length 237;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ASSTTNYT 8
```

```
US-10-424-599-207321

Query Match      80.5%; Score 33; DB 15; Length 118;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy      2 STTTNYT 8
Db      108 STSTNYT 114

RESULT 12
US-10-424-599-263645
; Sequence 263645, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 263645
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_80091C.1.pap
US-10-424-599-263645

Query Match      80.5%; Score 33; DB 15; Length 233;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ASSTTNYT 8
Db      113 ASSTTNYT 120

RESULT 13
US-10-425-114-40777
; Sequence 40777, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 40777
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3028-028-G1_FLI.pap
US-10-425-114-40777

Query Match      80.5%; Score 33; DB 15; Length 237;
Best Local Similarity 75.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ASSTTNYT 8
```

```
Db      117 ASSTNYT 124
||::|||
RESULT 14
US-10-767-701-46985
; Sequence 46985, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 46985
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C49_319.pep
US-10-767-701-46985
Query Match      80.5%; Score 33; DB 16; Length 476;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      3 TTTNYT 8
      |||||
Db      100 TTTNYT 105

RESULT 15
US-10-425-114-43218
; Sequence 43218, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 43218
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3067-008-A2_FLI.pep
US-10-425-114-43218
Query Match      80.5%; Score 33; DB 15; Length 483;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      3 TTTNYT 8
      |||||
Db      107 TTTNYT 112
```

Search completed: March 18, 2005, 19:15:20
Job time : 59.88 secs

THIS PAGE BLANK (05820)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 18:59:20 ; Search time 15.68 seconds
(without alignments)
49,090 Million cell updates/sec

Title: US-10-773-274A-1

Perfect score: 41

Sequence: 1 ASTTTNT 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_79.*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	855	1 VCLJAJ2	env polyprotein pr
2	37	90.2	504	2 T33404	hypothetical prote
3	36	87.8	283	2 G69212	conserved hypothe
4	35	85.4	551	2 H95283	probable seralysi
5	35	85.4	809	1 Q08E34	BBLF4 protein - hu
6	34	82.9	348	2 T11507	NADH2 dehydrogenas
7	34	82.9	759	2 S67164	probable membrane
8	33	80.5	635	2 S57114	capB protein - Clo
9	33	80.5	819	2 S15169	ferric-pseudobacti
10	33	80.5	1251	2 T21389	hypothetical prote
11	32	78.0	196	1 S50709	probable O-acetyl
12	32	78.0	335	2 AC0570	FimH protein precu
13	32	78.0	340	2 H88639	protein C34H4.1 li
14	32	78.0	393	2 T32533	hypothetical prote
15	32	78.0	529	2 T23181	hypothetical prote
16	32	78.0	552	2 T24690	hypothetical prote
17	32	78.0	592	2 E89772	hypothetical prote
18	32	78.0	791	2 T27473	hypothetical prote
19	32	78.0	928	1 VGBEMC	hypocyprotein B pre
20	32	78.0	1180	2 E86719	hypothetical prote
21	31	75.6	257	2 T03136	hypothetical prote
22	31	75.6	305	2 D97858	integrase/recombin
23	31	75.6	314	2 I68889	melanoma antigen M
24	31	75.6	500	2 F88921	protein F56E10.3 f
25	30	73.2	273	2 T19359	hypothetical prote
26	30	73.2	390	2 T40924	nuclear localizati
27	30	73.2	647	2 S48471	probable membrane
28	30	73.2	672	2 T30374	probable envelope
29	30	73.2	694	2 F97279	TPR-repeat-contain

30	30	73.2	722	2 T37970	probable G2-specif
31	30	73.2	917	2 F95884	probable sensory h
32	30	73.2	1576	2 S65774	homeotic protein H
33	30	73.2	1578	2 AD1512	peptidoglycan boun
34	30	73.2	1582	2 AC1153	adhesin homolog lm
35	30	73.2	2529	2 B64635	toxin-like outer m
36	29	70.7	34	2 D81044	hypothetical prote
37	29	70.7	86	2 S54912	ig heavy chain V r
38	29	70.7	110	2 G84629	hypothetical prote
39	29	70.7	149	2 E84629	hypothetical prote
40	29	70.7	150	2 T28297	hypothetical prote
41	29	70.7	214	2 S90125	hypothetical prote
42	29	70.7	247	2 E90493	hypothetical prote
43	29	70.7	354	2 B95884	probable sugar ABC
44	29	70.7	356	2 T37136	hypothetical prote
45	29	70.7	364	2 T47198	H+-exporting ATPas

ALIGNMENTS

RESULT 1

VCLJAJ2

env polyprotein precursor - human immunodeficiency virus type 1 (isolate ARV-2)
N;Alternate names: coat polyprotein
C;Species: human immunodeficiency virus type 1, HIV-1
A;Note: host Homo sapiens (man)
C;Date: 17-May-1985 #sequence_revision 17-May-1985 #text_change 09-Jul-2004
C;Accession: A03976
R;Sanchez-Pescador, R.; Power, M.D.; Barr, P.J.; Steimer, K.S.; Stempien, M.M.; Brown-Sh-N
Science 227, 484-492, 1985
A;Title: Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2).
A;Reference number: A04003; MUID:85090453; PMID:2578227
A;Accession: A03976
A;Molecule type: DNA
A;Residues: 1-855 <SAN>
A;Cross-references: UNIPROT:P03378; GB:K02007; NID:g328658; PIDN:AAB59882.1; PID:g328666
C;Genetics:
A;Gene: env
C;Superfamily: type E retrovirus env polyprotein
C;Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot
F;1-30/Domain: signal sequence #status predicted <SIG>
F;31-509/Product: exterior membrane glycoprotein #status predicted <EXT>
F;510-855/Product: transmembrane glycoprotein #status predicted <TM>
F;87,129,140,158,184,190,200,244,265,292,298,304,334,341,358,364,388,394,400,408,445,458
F;610,624,636,815/Binding site: carbohydrate (Aan) (covalent) #status predicted

Query Match 100.0%; Score 41; DB 1; Length 855;
Best Local Similarity 100.0%; Pred. NO. 2.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ASTTTNT 8
Db 185 ASTTTNT 192

RESULT 2

T33404

hypothetical protein H10E21.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T33404
R;Davidson, S.; Wohldmann, P.; Courtney, L.
submitted to the EMBL Data Library, July 1998
A;Description: The sequence of C. elegans cosmid H10E21.
A;Reference number: Z21339
A;Accession: T33404
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-504 <DAV>
A;Cross-references: UNIPROT:Q81TW8; EMBL:AF078783; PIDN:AAC26921.1; GSPDB:GN00021; CSSP:;
A;Experimental source: strain Bristol N2; clone H10E21
C;Genetics:

A;Gene: CBSP:H10E21.3
A;Map position: 3
A;Introns: 78/1; 168/3; 241/3; 384/1; 434/3

Query Match 90.2%; Score 37; DB 2; Length 504;
Best Local Similarity 87.5%; Pred. No. 10;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTTTNNT 8
|||||
DB 184 ASTTTNYS 191

RESULT 3
G69212
Conserved hypothetical protein MTH845 - Methanobacterium thermoautotrophicum (strain Delta H)
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: G69212
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadator, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functional genome
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: G69212
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-283 <MTH>
A;Cross-references: UNIPROT:O26933; GB:AE000861; GB:AE000666; NID:G2621930; PIDN:AAB8534
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH845

Query Match 87.8%; Score 36; DB 2; Length 283;
Best Local Similarity 87.5%; Pred. No. 8.9;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASTTTNNT 8
|||||
DB 53 ASTTTNNT 60

RESULT 4
H95263
Probable serralysin (EC 3.4.24.40) [imported] - Sinorhizobium meliloti (strain 1021) meg
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: H95263
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows, L.; Kalman, S.; Keating, D.H.; Palm, C.; Beck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.; Proc. Natl. Acad. Sci. U.S.A. 98, 9893-9898, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: H95263
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-551 <KUR>
A;Cross-references: UNIPROT:Q931C8; GB:AE006469; PIDN:AAK64674.1; PID:G14523072; GSPDB:C
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, H.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMA0034
A;Genome: plasmid
C;Keywords: hydrolase; metalloproteinase

Query Match 85.4%; Score 35; DB 2; Length 551;
Best Local Similarity 87.5%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASTTTNNT 8
|||||
DB 178 ASTTTGYT 185

RESULT 5
Q0BE34
BBLF4 protein - human herpesvirus 4 (strain B95-8)
C;Species: human herpesvirus 4, Epstein-Barr virus
C;Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 09-Jul-2004
C;Accession: F43043; A03776; S33026
R;Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
Mol. Biol. Med. 1, 21-45, 1983
A;Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
A;Reference number: A93065; MUID:85035713; PMID:6092825
A;Accession: F43043
A;Molecule type: DNA
A;Residues: 1-809 <BAN>
A;Cross-references: UNIPROT:P03214; EMBL:V01555; NID:G59074; PIDN:CAA24821.1; PID:G13348
R;Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H. Nature 310, 207-211, 1984
A;Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
A;Reference number: A03794; MUID:84270667; PMID:6087149
A;Contents: annotation; protein coding region
C;Superfamily: varicella-zoster virus gene 55 protein

Query Match 85.4%; Score 35; DB 1; Length 809;
Best Local Similarity 87.5%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASTTTNNT 8
|||||
DB 652 ASTTENVT 659

RESULT 6
T11507
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 2 - baboon mitochondrion (fragment)
C;Species: mitochondrion Papio hamadryas (baboon)
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T11507
R;Arnason, U.; Gullberg, A.; Janke, A.
J. Mol. Evol. 47, 718-727, 1998
A;Title: Molecular timing of primate divergences as estimated by two non-primate calibra
A;Reference number: Z17277; MUID:99065765; PMID:9847414
A;Accession: T11507
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-348 <ARN>
A;Cross-references: UNIPROT:Q92XY3; EMBL:Y18001; NID:G4049475; PIDN:CAA76995.1; PID:G404
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCI
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 2
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 82.9%; Score 34; DB 2; Length 348;
Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 STTTNNT 8
:|||||
DB 85 TTTTNT 91

RESULT 7
S67164
probable membrane protein YOR267c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein O5420
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence revision 12-Jul-1996 #text_change 16-Aug-2004
C:Accession: S67164; S67169; S72039
R:Jauniaux, J.C.; Poirrey, R.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67143
A:Accession: S67164
A:Molecule type: DNA
A:Residues: 1-759 <JAU>
A:Cross-references: UNIPROT:Q08732; EMBL:Z75175; NID:G1420599; PIDN:CAA99490.1; PID:e252
A:Experimental source: strain S288C
R:Cheret, G.; Sor, F.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67169
A:Accession: S67169
A:Molecule type: DNA
A:Residues: 1-367 <CHB>
A:Cross-references: EMBL:Z75175; MIPS:YOR267c
A:Experimental source: strain S288C
R:Cheret, G.; Bernardi, A.; Sor, F.
Yeast 12, 1059-1064, 1996
A:Title: DNA sequence analysis of the VPB1-SNF2 region on chromosome XV of Saccharomyces
A:Reference number: S72039; MUID:97051594; PMID:8896271
A:Accession: S72039
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-367 <CHW>
A:Cross-references: EMBL:X89633
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995
C:Genetics:
A:Map position: 15R
A:Cross-references: SGD:S0005793
A:Note: YOR267c
C:Superfamily: protein kinase homology
C:Keywords: ATP; transmembrane protein
F:213-496/Domain: protein kinase homology <KIN>
F:221-229/Region: protein kinase ATP-binding motif
F:400-416/Domain: transmembrane #status predicted <TMM>
F:628-656/Region: glutamine-rich

Query Match 82.9%; Score 34; DB 2; Length 759;
Best Local Similarity 85.7%; Pred. No. 60;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 STTTNYT 8
:|||||
Db 64 TTTTNYT 70

RESULT 8
S57714
capB protein - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S57714
R:Sanchez-Beato, A.; Garcia, J.
submitted to the EMBL Data Library, July 1995
A:Description: Molecular characterization of a family of choline-binding proteins of Clo
A:Reference number: S57714
A:Accession: S57714
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-635 <SAN>
A:Cross-references: UNIPROT:Q45820; EMBL:Z50008; NID:g895756; PIDN:CAA90303.1; PID:g8957
F:522-541/Domain: cpl repeat homology <CPI>
F:542-561/Domain: cpl repeat homology <CP2>
F:602-621/Domain: cpl repeat homology <CP3>

Query Match 80.5%; Score 33; DB 2; Length 635;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTTTNYT 8
:|||||
Db 376 TTTTNYT 381

RESULT 9
S15169
ferric-pseudobactin receptor precursor - Pseudomonas putida
C:Species: Pseudomonas putida
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C:Accession: S15169
R:Bitter, W.; Marugg, J.D.; de Weger, L.A.; Tommassen, J.; Weisbeek, P.J.
Mol. Microbiol. 5, 647-655, 1991
A:Title: The ferric-pseudobactin receptor Pupa of Pseudomonas putida WCS358: homology to
A:Reference number: S15169; MUID:91260449; PMID:1646376
A:Accession: S15169
A:Molecule type: DNA
A:Residues: 1-819 <BIT>
A:Cross-references: UNIPROT:P25184; EMBL:X56605; NID:g45722; PIDN:CAA39942.1; PID:g45723
C:Genetics:
A:Gene: pupA
C:Superfamily: ferric-pseudobactin receptor; tonB-dependent receptor amino-terminal homo
C:Keywords: membrane protein
F:1-47/Domain: signal sequence #status predicted <SIG>
F:48-819/Product: ferric-pseudobactin receptor #status predicted <MAT>
F:189-322/Domain: tonB-dependent receptor amino-terminal homology <TNN>
F:542-819/Domain: tonB-dependent receptor carboxyl-terminal homology <TNC>

Query Match 80.5%; Score 33; DB 2; Length 819;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TTTTNYT 8
:|||||
Db 396 TTTTNYT 401

RESULT 10
T21389
hypothetical protein F26C11.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T21389
R:Matthews, P.
submitted to the EMBL Data Library, December 1994
A:Reference number: Z19416
A:Accession: T21389
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1251 <WIL>
A:Cross-references: UNIPROT:Q09550; EMBL:Z47072; PIDN:CAA87369.1; GSPDB:GN00020; CESP:F.2
A:Experimental source: clone F26C11
C:Genetics:
A:Gene: CESP:F26C11.3
A:Map position: 2
A:Introns: 24/1; 111/1; 208/1; 272/1; 380/1; 394/3; 485/3; 586/3; 630/3; 669/3; 713/3; 7

Query Match 80.5%; Score 33; DB 2; Length 1251;
Best Local Similarity 85.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 STTTNYT 8
:|||||
Db 393 STTTNYT 399

RESULT 11
S50709
Probable O-acetyltransferase (EC 2.3.1.-) YJL218w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein HRA196; hypothetical protein J0224
C:Species: Saccharomyces cerevisiae
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S50709; S57008; S45154

R;Vandenbol, M.; Durand, P.; Bolle, P.A.; Dion, C.; Portetelie, D.; Hilger, F.
Yeast 10, 1657-1662, 1994
A>Title: Sequence analysis of a 40.2 kb DNA fragment located near the left telomere of Y
A;Reference number: S50701; MUID:95242842; PMID:7725802
A;Accession: S50709
A>Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-196 <VAN>
A;Cross-references: UNIPROT:P40892; EMBL:Z34098; NID:G496934; PIDN:CAA8392.1; PID:G4969
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994
R;Vandenbol, M.; Durand, P.; Portetelie, D.; Hilger, F.
submitted to the Protein Sequence Database, September 1995
A;Reference number: S57008
A;Accession: S57008
A;Molecule type: DNA
A;Residues: 1-196 <VAV>
A;Cross-references: EMBL:Z49493; NID:G1015607; PIDN:CAA89515.1; PID:G1015608; GSPDB:GN00
C;Genetics:
A;Gene: MIPS:YJL218w
A;Cross-references: SGD:S0003754
A;Map position: 10L
C;Superfamily: galactoside acetyltransferase
C;Keywords: acyltransferase

Query Match 78.0%; Score 32; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 STTTNY 7
| | | | |
Db 189 STTTNY 194

RESULT 12
AC0570
FlnH protein precursor [imported] - Salmonella enterica subsp. enterica serovar Typhi (S
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AC0570
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AC0570
A>Status: preliminary
A;Molecule type: DNA
A;Residues: 1-335 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05029.1; PID:G16501811; GSPDB:GN00176
C;Genetics:
A;Gene: flmH

Query Match 78.0%; Score 32; DB 2; Length 335;
Best Local Similarity 75.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ASTTTNYT 8
| | | | |
Db 71 AGTTVNYT 78

RESULT 13
H8639
protein C34H4.1 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: H8639
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog

A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: H8639
A>Status: preliminary
A;Molecule type: DNA
A;Residues: 1-340 <STO>
A;Cross-references: UNIPROT:O45144; GB:Chr_IV; PIDN:AAC04390.1; PID:G2911817; GSPDB:GN00
C;Genetics:
A;Gene: C34H4.1
A;Map position: 4

Query Match 78.0%; Score 32; DB 2; Length 340;
Best Local Similarity 75.0%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASTTTNYT 8
| | | | |
Db 256 AATTGTGT 263

RESULT 14
T32533
hypothetical protein CD4.9 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T32533
R;Du, Z.; Scheet, P.
submitted to the EMBL Data Library, December 1997
A;Description: The sequence of C. elegans cosmid CD4.
A;Reference number: Z21185
A;Accession: T32533
A>Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-393 <DUZ>
A;Cross-references: UNIPROT:O44163; EMBL:AF036694; PIDN:AAB88347.1; GSPDB:GN00022; CESP:
A;Experimental source: strain Bristol N2; clone CD4
C;Genetics:
A;Gene: CESP:CD4.9
A;Map position: 4
A;Introns: 21/2; 71/2; 100/1; 195/3; 209/3; 319/2; 379/1

Query Match 78.0%; Score 32; DB 2; Length 393;
Best Local Similarity 75.0%; Pred. No. 77;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASTTTNYT 8
| | | | |
Db 145 ATTTTNYT 152

RESULT 15
T33181
hypothetical protein K01D12.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T33181
R;Dobson, R.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z19703
A;Accession: T33181
A>Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-529 <WIL>
A;Cross-references: UNIPROT:Q21093; EMBL:Z75543; PIDN:CAA99870.1; GSPDB:GN00023; CESP:K
A;Experimental source: clone K01D12
C;Genetics:
A;Gene: CESP:K01D12.2
A;Map position: 5
A;Introns: 152/2
C;Superfamily: Caenorhabditis elegans hypothetical protein K01D12.2

Query Match 78.0%; Score 32; DB 2; Length 529;

Best Local Similarity 85.7%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 STTNYT 8
|||:|
Db 478 STTHYT 484

Search completed: March 18, 2005, 19:11:05
Job time : 19.68 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 18:50:09 ; Search time 73.28 Seconds
(without alignments)
55.904 Million cell updates/sec

Title: US-10-773-274A-1
Perfect score: 41
Sequence: 1 ASTTTNYT 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	100.0	853	2 Q80161	Q80161 human immun
2	41	100.0	855	1 ENV_HV1A2	P03378 human immun
3	37	90.2	350	2 Q9BJK7	Q9BJK7 caenorhabdi
4	37	90.2	486	2 Q76668	O76668 caenorhabdi
5	37	90.2	504	2 Q8ITW8	Q8ITW8 caenorhabdi
6	37	90.2	860	2 Q9YP43	Q9YP43 human immun
7	37	90.2	860	2 Q9YP50	Q9YP50 human immun
8	37	90.2	861	2 Q9WJV5	Q9WJV5 human immun
9	37	90.2	864	2 Q9YPA8	Q9YPA8 human immun
10	37	90.2	900	2 Q8KQ11	Q8KQ11 burkholderi
11	37	90.2	900	2 Q8KQJ8	Q8KQJ8 burkholderi
12	36	87.8	283	2 Q26933	Q26933 methanobact
13	35	85.4	212	2 Q81RD7	Q81RD7 bacillus an
14	35	85.4	551	2 Q931C8	Q931C8 rhizobium m
15	35	85.4	809	1 HELI_EBV	P03214 epstein-bar
16	35	85.4	809	2 Q8UZG7	Q8UZG7 cercopithec
17	35	85.4	809	2 Q777D7	Q777D7 human herpe
18	34	82.9	348	1 NU2M_PAPHA	Q9ZXY3 papio hamad
19	34	82.9	360	2 Q9HED8	Q9HED8 neurospora
20	34	82.9	368	2 Q07224	Q07224 saccharomyc
21	34	82.9	704	2 Q8JKL0	Q8JKL0 heliothis z
22	34	82.9	759	2 Q08732	Q08732 saccharomyc
23	33	80.5	65	2 Q6ZIX6	Q6ZIX6 oryza sativ
24	33	80.5	127	2 Q6QWE0	Q6QWE0 rattus norv
25	33	80.5	154	2 Q89M41	Q89M41 bradyrhizob
26	33	80.5	258	2 Q89501	Q89501 human immun
27	33	80.5	284	2 Q8DDT9	Q8DDT9 xenopus lae
28	33	80.5	320	2 Q9VMB1	Q9VMB1 drosophila
29	33	80.5	352	2 Q8EGF5	Q8EGF5 shewanella
30	33	80.5	378	2 Q8EVB1	Q8EVB1 mycoplasma
31	33	80.5	478	2 Q717T5	Q717T5 drosophila

32 33 80.5 559 2 Q842B3
33 33 80.5 589 2 Q985N8
34 33 80.5 619 2 Q7RD05
35 33 80.5 635 2 Q45820
36 33 80.5 668 2 Q813R2
37 33 80.5 682 2 Q6FIM0
38 33 80.5 819 1 PUPA_PSEFU
39 33 80.5 868 2 Q41575
40 33 80.5 1240 1 YQU3_CABEL
41 33 80.5 1754 2 Q861F3
42 32 78.0 80 2 Q80LS3
43 32 78.0 157 2 Q15620
44 32 78.0 167 2 Q9D9H2
45 32 78.0 196 1 YJV8_YEAST

ALIGNMENTS

RESULT 1
Q80161 PRELIMINARY; PRT; 853 AA.
AC Q80161
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope polyprotein.
GN Name=env;
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92046357; PubMed=1658383;
RA Cheng-Mayer C., Shioda T., Levy J.A.;
RT "Host range, replicative, and cytopathic properties of human
RT immunodeficiency virus type 1 are determined by very few amino acid
RT changes in tat and gp120.";
RL J. Virol. 65:6931-6941(1991).
DR EMBL; L07422; AAA80324.1; -.
DR HSPG; P04578; IDLB.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR InterPro; IPR011010; DNA_brk_join_enz.
DR InterPro; IPR000328; Env_Gp41.
DR InterPro; IPR000777; Gp120.
DR Pfam; PF00516; Gp120; 1.
DR Pfam; PF00517; Gp41; 1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
SQ SEQUENCE 853 AA; 97056 MW; 2181503CFD14789 CRC64;
Query Match 100.0%; Score 41; DB 2; Length 853;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ASTTTNYT 8
Db 184 ASTTTNYT 191
RESULT 2
ENV_HV1A2 STANDARD; PRT; 855 AA.
AC P03378;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUN-2004 (Rel. 44, Last annotation update)
DE Envelope polyprotein Gp160 precursor [Contains: Exterior membrane
DE glycoprotein (Gp120); Transmembrane glycoprotein (Gp41)].
GN Name=ENV;

OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11685;
RN SEQUENCE FROM N.A.
RP MEDLINE=8509453; PubMed=2578227;
RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,
RA Stempien M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,
RA Levy J.A., Dina D., Luciw P.A.;
RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
(ARV-2)".
RL Science 227:484-492 (1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02007; AAB59882.1; -;
DR PIR; A03976; VCLJJA2.
DR HSSP; P04578; 1DLB.
DR HIV; K02007; ENVSSF2.
DR InterPro; IPR011010; DNA_brk_join_enz.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Signal; Transmembrane.
FT SIGNAL 1
FT CHAIN 30 509 Exterior membrane glycoprotein.
FT CHAIN 510 855 Transmembrane glycoprotein.
FT DISULFID 53 73 By similarity.
FT DISULFID 118 208 By similarity.
FT DISULFID 125 199 By similarity.
FT DISULFID 130 155 By similarity.
FT DISULFID 221 250 By similarity.
FT DISULFID 231 242 By similarity.
FT DISULFID 299 333 By similarity.
FT DISULFID 380 442 By similarity.
FT DISULFID 387 415 By similarity.
FT CARBOHYD 87 87 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 129 129 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 140 140 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 154 154 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 158 158 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 184 184 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 190 190 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 200 200 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 233 233 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 244 244 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 265 265 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 279 279 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 292 292 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 298 298 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 304 304 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 334 334 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 341 341 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 358 358 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 364 364 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 388 388 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 394 394 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 400 400 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 408 408 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 445 445 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 458 458 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 461 461 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 610 610 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 615 615 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 624 624 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 636 636 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 815 815 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 855 AA; 97438 MW; A3BC20573AAC41A2 CRC64;
Query Match 100.0%; Score 41; DB 1; Length 855;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ASSTTNYT 8
Db 185 ASITTYNT 192
RESULT 3
Q9BJK7 PRELIMINARY; PRT; 350 AA.
AC Q9BJK7
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nuclear receptor NHR-80 (Fragment).
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Robinson-Rechavi M., Maina C.V., Gissendanner C., Laudet V.,
RA Sluder A.E.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF332209; AAK17980.1; -;
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR00536; Hrmn_recept_lig.
DR InterPro; IPR008946; Str_ncl_receptor.
DR Pfam; PF00104; Hormone_recep; 1.
DR SMART; SM00430; HOL1; 1.
KW Receptor.
FT NON_TER 1
SQ SEQUENCE 350 AA; 40181 MW; DB3BB365E1427349 CRC64;
Query Match 90.2%; Score 37; DB 2; Length 350;
Best Local Similarity 87.5%; Pred. No. 21;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ASSTTNYT 8
Db 30 ASITTYNTS 37
RESULT 4
O76668 PRELIMINARY; PRT; 486 AA.
AC O76668
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Nuclear hormone receptor family protein 80, isoform a (Nuclear
DE receptor NHR-80).
GN Name=nhr-80; ORFNames=H10E21.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";

RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Davidson S., Wohldmann P., Courtney L.;
RT "The sequence of C. elegans fosmid H10E21.1";
RN Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Robinson-Rechavi M., Maina C.V., Gissendanner C.R., Laudet V.,
RA Sluder A.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR EMBL; AF078783; AAK82901.1; -.
DR EMBL; AY204179; AAO39183.1; -.
DR HSSP; P34021; 1R00.
DR WormBase; WBGene0003670; H10E21.
DR WormPep; H10E21.3a; CE28733.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000536; Hmon_recept_lig.
DR InterPro; IPR008946; Str_ncl_receptor.
DR InterPro; IPR000324; VitD_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; Hormone_recep; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00047; STROIDFINGER.
DR PRINTS; PR00350; VITAMINDR.
DR ProDom; PR000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; Znf_C4; 1.
KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
SQ SEQUENCE 486 AA; 55110 MW; 3424B4DD5C35BC64 CRC64;

Query Match 90.2%; Score 37; DB 2; Length 486;
Best Local Similarity 87.5%; Pred. NO. 30;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTTTNVT 8
Db |||||: 173
166 ASTTINYS 173

RESULT 5
Q81TW8 PRELIMINARY; PRT; 504 AA.
ID Q81TW8
AC Q81TW8
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nuclear hormone receptor family protein 80, isoform b.
GN Name=nhr-80; ORFNames=H10E21.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;

RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Davidson S., Wohldmann P., Courtney L.;
RT "The sequence of C. elegans fosmid H10E21.1";
RN Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family.
DR EMBL; AF078783; AAK83404.1; -.
DR PIR; T33404; T33404.
DR HSSP; P34021; 1R00.
DR WormBase; WBGene0003670; H10E21.3.
DR WormPep; H10E21.3b; CE19488.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003707; F:steroid hormone receptor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000536; Hmon_recept_lig.
DR InterPro; IPR008946; Str_ncl_receptor.
DR InterPro; IPR000324; VitD_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; Hormone_recep; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00047; STROIDFINGER.
DR PRINTS; PR00350; VITAMINDR.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; Znf_C4; 1.
KW DNA-binding; Metal-binding; Nuclear protein; Receptor; Transcription;
KW Transcription regulation; Zinc; Zinc-finger.
SQ SEQUENCE 504 AA; 57345 MW; 06B95A963B4E0BB4 CRC64;

Query Match 90.2%; Score 37; DB 2; Length 504;
Best Local Similarity 87.5%; Pred. NO. 32;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ASTTTNVT 8
Db |||||: 191
184 ASTTINYS 191

RESULT 6
Q9YP43 PRELIMINARY; PRT; 860 AA.
ID Q9YP43
AC Q9YP43
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope polypeptide.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21535983; PubMed=11679152; DOI=10.1089/08982201753197060;
RA Fang G., Moran H., Chappey C., Rowland-Jones S., Visosky A.,
RA Chen C.H., Moran T., Townsend L., Murray M., Weiser B.;
RT "Analysis of transition from long-term nonprogressive to progressive
RT infection identifies sequences that may attenuate HIV type 1.";

```
RL AIDS Res. Hum. Retroviruses 17:1395-1404(2001).
DR EMBL; U69588; AAD10907.1; -.
DR HSSP; P04578; 1DLB.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0000328; Env GP41.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
SQ SEQUENCE 860 AA; 97738 MW; 1FEB24AADF7F4216 CRC64;

Query Match 90.2%; Score 37; DB 2; Length 860;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 STTTNYT 8
Db 192 STTTNYT 198

RESULT 7
Q9YP50 PRELIMINARY; PRT; 860 AA.
ID Q9YP50
AC Q9YP50;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope polyprotein.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21535983; PubMed=11679152; DOI=10.1089/089922201753197060;
RA Fang G., Burger H., Chappey C., Rowland-Jones S., Visosky A.,
RA Chen C.H., Moran T., Townsend L., Murray M., Weiser B.;
RT "Analysis of transition from long-term nonprogressive to progressive
RT infection identifies sequences that may attenuate HIV type 1.";
RL AIDS Res. Hum. Retroviruses 17:1395-1404(2001).
DR EMBL; U69585; AAD10883.1; -.
DR HSSP; P04578; 1DLB.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
SQ SEQUENCE 860 AA; 97661 MW; A5AC60C59B990BC CRC64;

Query Match 90.2%; Score 37; DB 2; Length 860;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 STTTNYT 8
Db 192 STTTNYT 198

RESULT 8
Q9WJV5 PRELIMINARY; PRT; 861 AA.
ID Q9WJV5
AC Q9WJV5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope polyprotein.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21535983; PubMed=11679152; DOI=10.1089/089922201753197060;
RA Fang G., Burger H., Chappey C., Rowland-Jones S., Visosky A.,
RA Chen C.H., Moran T., Townsend L., Murray M., Weiser B.;
RT "Analysis of transition from long-term nonprogressive to progressive
RT infection identifies sequences that may attenuate HIV type 1.";
RL AIDS Res. Hum. Retroviruses 17:1395-1404(2001).
DR EMBL; U69585; AAD10883.1; -.
DR HSSP; P04578; 1DLB.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
SQ SEQUENCE 860 AA; 97661 MW; A5AC60C59B990BC CRC64;

Query Match 90.2%; Score 37; DB 2; Length 860;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 STTTNYT 8
Db 192 STTTNYT 198

RESULT 9
Q9YP48 PRELIMINARY; PRT; 864 AA.
ID Q9YP48
AC Q9YP48;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope polyprotein.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21535983; PubMed=11679152; DOI=10.1089/089922201753197060;
RA Fang G., Burger H., Chappey C., Rowland-Jones S., Visosky A.,
RA Chen C.H., Moran T., Townsend L., Murray M., Weiser B.;
RT "Analysis of transition from long-term nonprogressive to progressive
RT infection identifies sequences that may attenuate HIV type 1.";
RL AIDS Res. Hum. Retroviruses 17:1395-1404(2001).
DR EMBL; U69586; AAD10891.1; -.
DR HSSP; P04578; 1DLB.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
SQ SEQUENCE 864 AA; 98167 MW; 2A0E7AB2A89662F2 CRC64;

Query Match 90.2%; Score 37; DB 2; Length 864;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 STTTNYT 8
```

```
DE Envelope polyprotein.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21535983; PubMed=11679152; DOI=10.1089/089922201753197060;
RA Fang G., Burger H., Chappey C., Rowland-Jones S., Visosky A.,
RA Chen C.H., Moran T., Townsend L., Murray M., Weiser B.;
RT "Analysis of transition from long-term nonprogressive to progressive
RT infection identifies sequences that may attenuate HIV type 1.";
RL AIDS Res. Hum. Retroviruses 17:1395-1404(2001).
DR EMBL; U69587; AAD10899.1; -.
DR HSSP; P04578; 1DLB.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
SQ SEQUENCE 861 AA; 97993 MW; 3E3EFD977CC18EA CRC64;

Query Match 90.2%; Score 37; DB 2; Length 861;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 STTTNYT 8
Db 193 STTTNYT 199

RESULT 9
Q9YP48 PRELIMINARY; PRT; 864 AA.
ID Q9YP48
AC Q9YP48;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope polyprotein.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21535983; PubMed=11679152; DOI=10.1089/089922201753197060;
RA Fang G., Burger H., Chappey C., Rowland-Jones S., Visosky A.,
RA Chen C.H., Moran T., Townsend L., Murray M., Weiser B.;
RT "Analysis of transition from long-term nonprogressive to progressive
RT infection identifies sequences that may attenuate HIV type 1.";
RL AIDS Res. Hum. Retroviruses 17:1395-1404(2001).
DR EMBL; U69586; AAD10891.1; -.
DR HSSP; P04578; 1DLB.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
KW Transmembrane.
SQ SEQUENCE 864 AA; 98167 MW; 2A0E7AB2A89662F2 CRC64;

Query Match 90.2%; Score 37; DB 2; Length 864;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 STTTNYT 8
```



```

Db      192 STTTNYT 198
|||||
RESULT 10
O8KQJ11 PRELIMINARY; PRT; 900 AA.
AC O8KQJ11;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE Putative outer membrane usher.
GN Name=cblC;
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BC7;
RX PubMed=14761995; DOI=10.1128/JB.186.4.1009-1020.2004;
RA Tomich M., Mohr C.D.;
RT "Transcriptional and posttranscriptional control of cable pilus gene
expression in Burkholderia cenocepacia.";
RL J. Bacteriol. 186:1009-1020(2004).
DR EMBL; AY114293; AAM56039.1; -.
SQ SEQUENCE 900 AA; 96322 MW; D6FAFA1989E1392F CRC64;

Query Match 90.2%; Score 37; DB 2; Length 900;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 STTTNYT 8
|||||
Db 599 STTTNYT 605

RESULT 11
O8KQJ8 PRELIMINARY; PRT; 900 AA.
AC O8KQJ8;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE Putative usher protein.
GN Name=cblC;
OS Burkholderia cepacia (Pseudomonas cepacia).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22573395; PubMed=12686638; DOI=10.1099/mic.0.26176-0;
RA Sajjan U.S., Xie H., Lefebvre M.D., Valvano M.A., Forstner J.F.;
RT "Identification and molecular analysis of cable pilus biosynthesis
genes in Burkholderia cepacia.";
RL Microbiology 149:961-971(2003).
DR EMBL; AY082893; AAL92875.1; -.
SQ SEQUENCE 900 AA; 96416 MW; ASC437197365C2D9 CRC64;

Query Match 90.2%; Score 37; DB 2; Length 900;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 STTTNYT 8
|||||
Db 599 STTTNYT 605

RESULT 12
O26933 PRELIMINARY; PRT; 283 AA.
ID O26933
AC O26933;

```

```

DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Conserved protein.
GN OrderedLocusNames=MT845;
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lamm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
deltah: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
DR EMBL; AE000861; AAB85343.1; -.
DR PIR; G69212; G69212.
DR InterPro; IPR002931; Trnsglucanase like.
DR Pfam; PF01841; Transglut_core; 1.
DR SMART; SM00460; TGC; 1.
KW Complete proteome.
SQ SEQUENCE 283 AA; 32015 MW; DB9045876AB76976 CRC64;

Query Match 87.8%; Score 36; DB 2; Length 283;
Best Local Similarity 87.5%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ASTTTNYT 8
|||||
Db 53 ASTTTNYT 60

RESULT 13
O81RD7 PRELIMINARY; PRT; 212 AA.
ID O81RD7; O6HZL1; O6KTK1;
AC O81RD7; O6HZL1; O6KTK1;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BA2112, BAS1964, GBAA2112;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Ames / isolate Porton;
RX MEDLINE=22608414; PubMed=12731629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtapple E.K., Ostad O.A., Helgason E., Ristone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Niemman W.C.,
RA Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,
RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstoe A.-B., Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
closely related bacteria.";
RL Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;

```

Query Match 85.4%; Score 35; DB 2; Length 551;
Best Local Similarity 87.5%; Pred. No. 91;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RA Ravel J., Raeko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
RA "Bacillus anthracis comparative genomics."
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sterner;
RA Brettin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Kelm P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Rice H.,
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017030; AAP25996.1; -
DR EMBL; AE017334; AAT31230.1; -
DR EMBL; AE017225; AAT54278.1; -
DR TIGR; BA2112; -
DR TIGR; GBA2112; -
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 212 AA; 24724 MW; F557E29021107126 CRC64;

QY 1 ASTTNYT 8
Db 178 ASTTGYT 185

RESULT 15
HELI EBV STANDARD; PRT; 809 AA.
ID HEBV
AC P03214;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Probable helicase.
GN Name=BBLF4;
OS Epstein-Barr virus (strain B95-8) (HBV-4) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
OC NCBI_TaxID=10377;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Seguin C.,
RA Tuffnell P.S., Barrell B.G.;
RA "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RT Nature 310:207-211 (1984).
RL Nature 310:207-211 (1984).
CC -I- FUNCTION: This protein may be an helicase and is required for
CC replication of viral DNA.
CC -I- SIMILARITY: Belongs to the herpesviruses helicase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; V01555; CAA24821.1; -
DR PIR; F43043; Q0BE34.
DR InterPro; IPR003840; Herpes helicase.
DR Pfam; PF02689; Herpes Helicase; 1.
KW ATP-binding; DNA replication; Early protein; Helicase.
FT NP_BIND 72 79
SQ SEQUENCE 809 AA; 89853 MW; 434AA6EDAC01CC50 CRC64;

QY 1 ASTTNYT 8
Db 652 ASTTNYT 659

Query Match 85.4%; Score 35; DB 1; Length 809;
Best Local Similarity 87.5%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASTTNYT 8
Db 652 ASTTNYT 659

Search completed: March 18, 2005, 19:06:06
Job time : 77.28 secs

Query Match 85.4%; Score 35; DB 2; Length 212;
Best Local Similarity 87.5%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ASTTNYT 8
Db 49 ASTTNYT 56

RESULT 14
Q931C8 PRELIMINARY; PRT; 551 AA.
ID Q931C8;
AC Q931C8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Possible protease (EC 3.4.24.40).
GN ORFNames=SmA0034;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymA.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OC NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432; DOI=10.1073/pnas.161294798;
RA Barnett M.J., Fisher R.P., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federsteel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymA megaplasmid."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888 (2001).
RL EMBL; AE007196; AAK64674.1; -
DR PIR; H95263; H95263.
DR HSSP; O69771; 10M8.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0008237; F:metalloproteinase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001343; Hemlyan_Ca_bind.
DR InterPro; IPR006026; Peptidase_M.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR011049; Serralyen_like_C.
DR Pfam; PF00353; HemolysinCabin; 6.
DR Pfam; PF00413; HemolysinCabin; 6.
DR PRINTS; PR00313; CAENDNGRPT.
DR SMART; SM00235; ZnMc; 1.
DR PROSITE; PS00330; HEMOLYSIN_CALCIUM; 2.
KW Complete proteome; Hydrolase; Plasmid; Protease.
SQ SEQUENCE 551 AA; 57726 MW; DC5907F6F027D056 CRC64;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 18:51:00 ; Search time 161.84 Seconds
(without alignments)
40.626 Million cell updates/sec

Title: US-10-773-274A-2

Perfect score: 97

Sequence: 1 PERDISNVFSPDGKPC 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	1255	8	Adi95327 OSPF-rela
2	55	56.7	332	7	Abp63121 Mycelioph
3	55	56.7	332	8	Adp84784 Corynacu
4	55	56.7	332	8	Adp84795 Mutant C.
5	55	56.7	350	2	AAW27063 Mycelioph
6	51	52.6	846	7	Adb70258 C. neofo
7	50	51.5	655	8	AdS41636 Bacterial
8	50	51.5	1136	8	AdS30013 Bacterial
9	48	49.5	1155	8	AdS30836 Bacterial
10	47.5	49.0	699	8	AdS23978 Bacterial
11	47	48.5	134	3	AAy75490
12	47	48.5	256	7	ABO65040 Klebsiell
13	47	48.5	590	6	ABD12353 Alloiooc
14	47	48.5	590	8	AdJ27161 Alloiooc
15	47	48.5	626	8	AdS41817 Bacterial
16	47	48.5	642	2	AAy01541 Alpha(2-3
17	46	47.4	120	5	AdH32823 Yeast smo
18	46	47.4	508	8	AdN26478 Bacterial
19	46	47.4	613	8	AdS22148 Bacterial
20	46	47.4	756	8	AdS30912 Bacterial
21	45	46.4	207	8	AdN23184 Bacterial
22	45	46.4	275	8	AdS82240 Human pro
23	45	46.4	288	2	AAr22996 Yeast pro
24	45	46.4	288	4	ABG09463 Novel hum
25	45	46.4	288	6	ABR53144 Protein s

26	45	46.4	288	7	ADK63118	Adk63118 Disease t
27	45	46.4	288	8	ADS44003	AdS44003 Bacterial
28	45	46.4	377	8	ADH78730	Adh78730 Tobacco A
29	45	46.4	596	4	ABG09464	Abg09464 Novel hum
30	45	46.4	1118	8	ADG29812	Adg29812 Bacterial
31	44.5	45.9	117	3	AAG43935	Aag43935 Zea may
32	44.5	45.9	296	4	AAm25263	Aam25263 Human pro
33	44.5	45.9	414	8	ABO84990	AbO84990 Murine ca
34	44.5	45.9	451	8	ABO84991	AbO84991 Human can
35	44	45.4	27	8	ADJ78138	Adj78138 Peptide S
36	44	45.4	29	5	ABG96085	Abg96085 Cysteine-
37	44	45.4	108	4	AAO04124	Aao04124 Human pol
38	44	45.4	167	5	ABP06586	Abp06586 Human ORF
39	44	45.4	210	2	RAY73870	Aay73870 Human pro
40	44	45.4	395	8	ADN24340	Adn24340 Bacterial
41	44	45.4	548	7	ABO72047	AbO72047 Pseudomon
42	44	45.4	594	6	ABR53208	AbR53208 Protein s
43	44	45.4	594	7	ADK63748	Adk63748 Disease t
44	44	45.4	613	8	ADN19303	Adn19303 Bacterial
45	44	45.4	613	8	ADN18891	Adn18891 Bacterial

ALIGNMENTS

RESULT 1

ADI95327
ID ADI95327 standard; protein; 1255 AA.

XX AC ADI95327;

DT 04-NOV-2004 (first entry)

XX OS OSFP-related SARS coronavirus Frankfurt 1 spike protein S.

XX KW immune response; overlapping synthetic peptide formulation; OSFP;
immunostimulant; virucide; antibacterial; antiparasitic; cytostatic;
vaccine; viral; bacterial; parasitic infection; prion disease;
neoplastic; toxin; spike.

XX OS SARS coronavirus Frankfurt 1.

XX PN WO2004002415-A2.

XX PD 08-JAN-2004.

XX PF 27-JUN-2003; 2003WO-US020322.

XX PR 27-JUN-2002; 2002US-0392718P.

XX PA (DAND) DANA FARBER CANCER INST INC.

XX FI Ruprecht RM, Jiang S;

XX DR WPI, 2004-082868/08.

XX PT Modulating an immune response, useful for treating immune disorders, e.g.
viral, bacterial and parasitic infections, prion diseases, or neoplastic
diseases, administering to a subject an overlapping synthetic peptide
formulation.

XX PS Claim 13; SEQ ID NO 235; 175pp; English.

XX CC The invention relates to a novel method for modulating an immune response
comprising administering to a subject an overlapping synthetic peptide
formulation (OSPF) which comprises a combination of single chain peptides
corresponding to the amino acid sequence of a protein of interest. The
method of the invention has immunostimulant, virucide, antibacterial,
antiparasitic and cytostatic applications and may be useful during
vaccine production and for treating immune disorders including viral,
bacterial and parasitic infections, prion diseases, neoplastic diseases,
as well as providing protection against toxins. The current sequence is
that of the OSFP-related SARS coronavirus Frankfurt 1 spike protein S of

CC the invention.
 XX Sequence 1255 AA;
 SQ

Query Match 100.0%; Score 97; DB 8; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 7e-07;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FERDISNVFSPDGKPC 17
 |||||
 DB 451 FERDISNVFSPDGKPC 467

RESULT 2
 ABR63121
 ID ABR63121 standard; protein; 332 AA.
 XX
 AC ABR63121;
 XX 18-DEC-2003 (first entry)
 DT Myceliophthora thermophila galactanase.
 DE
 XX Galactanase; thermostable; enzyme; feed additive.
 KW
 XX Myceliophthora thermophila.
 OS
 XX WO2003062409-A2.
 XX 31-JUL-2003.
 PD
 XX 23-JAN-2003; 2003WO-DK000039.
 PF
 XX 25-JAN-2002; 2002DK-00000130.
 PR
 XX (HOFF) ROCHE VITAMINS AG.
 PA
 XX Wu W, Pettersson D, Fuglsang CC;
 PI WPI; 2003-731382/69.
 DR
 XX Composition useful as an animal feed additive comprises at least two
 PT thermostable enzymes selected from endoglucanase, xylanase, phytase,
 PT protease, galactanase, mannanase, dextranase and alpha-galactosidase.
 XX Disclosure; Page 63-64; Opp; English.
 PS
 XX The present sequence is the protein sequence of a thermostable
 CC galactanase (without the signal peptide) of Myceliophthora thermophila.
 CC Preferred variants of the galactanase can be used in a claimed
 CC composition of the invention, which comprises at least 2 thermostable
 CC enzymes selected from an endoglucanase, xylanase, phytase, protease,
 CC galactanase, mannanase, dextranase and alpha-galactosidase. The
 CC composition is useful for improving the nutritional value of animal
 CC feeds, especially those containing soya, wheat, barley, oats and/or rye
 XX
 SQ Sequence 332 AA;

Query Match 56.7%; Score 55; DB 7; Length 332;
 Best Local Similarity 53.3%; Pred. No. 1.7;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 FERDISNVFSPDGK 15
 |||||
 DB 259 FPSDVKNIPSPGQ 273

RESULT 3
 ADP84784
 ID ADP84784 standard; protein; 332 AA.
 XX
 AC ADP84784;
 XX

23-SEP-2004 (first entry)
 Corynascus heterothallicus glycoside hydrolase family 53 galactanase.
 glycoside hydrolase family 53 galactanase; GH family 53 galactanase;
 dairy industry; galacto-oligosaccharide preparation; lactose hydrolysis;
 enzyme.
 OS Corynascus heterothallicus.
 XX WO2004056988-A2.
 PN 08-JUL-2004.
 PD 11-DEC-2003; 2003WO-DK000851.
 PF 20-DEC-2002; 2002DK-00001968.
 XX 08-APR-2003; 2003DK-00000537.
 PR (NOVO) NOVOZYMES AS.
 PA De Maria L, Svendsen A, Borchert TV, Christensen LLH, Larsen S;
 XX Ryttersgaard C;
 PI WPI; 2004-507720/48.
 DR Novel variant of parent glycoside hydrolase family 53 galactanase, useful
 XX in dairy industry, to prepare galacto-oligosaccharide and/or for
 PT hydrolysis of lactose.
 FT Claim 19; SEQ ID NO 1; 262pp; English.
 XX The invention comprises variants of a glycoside hydrolase (GH) family 53
 CC galactanase. The GH family 53 galactanase variants of the invention are
 CC useful in the dairy industry for the preparation of galacto-
 CC oligosaccharide and the hydrolysis of lactose. The present amino acid
 CC sequence represents a wild type Corynascus heterothallicus GH family 53
 CC galactanase of the invention.
 XX
 SQ Sequence 332 AA;

Query Match 56.7%; Score 55; DB 8; Length 332;
 Best Local Similarity 53.3%; Pred. No. 1.7;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 FERDISNVFSPDGK 15
 |||||
 DB 259 FPSDVKNIPSPGQ 273

RESULT 4
 ADP84795
 ID ADP84795 standard; protein; 332 AA.
 XX
 AC ADP84795;
 XX 23-SEP-2004 (first entry)
 DT Mutant C. heterothallicus glycoside hydrolase family 53 galactanase.
 XX glycoside hydrolase family 53 galactanase; GH family 53 galactanase;
 KW dairy industry; galacto-oligosaccharide preparation; lactose hydrolysis;
 KW enzyme; mutant; mutein.
 XX Corynascus heterothallicus.
 OS Synthetic.
 XX Key Location/Qualifiers
 FT Misc-difference 90 /note= "Wild type Ala replaced by Ser"
 FT Misc-difference 91 /note= "Wild type His replaced by Asp"
 FT
 XX

```

PN WO2004056988-A2.
XX
XX
PD 08-JUL-2004.
XX
XX PF 11-DEC-2003; 2003WO-DK0000851.
XX
XX PR 20-DEC-2002; 2002DK-00001968.
XX
XX PR 08-APR-2003; 2003DK-00000537.
XX
XX PA (NOVO ) NOVOZYMES AS.
XX
XX PI De Maria L, Svendsen A, Borchert TV, Christensen LLH, Larsen S;
XX
XX PI Ryttersgaard C;
XX
XX DR WPI; 2004-507720/48.
XX
XX
XX PT Novel variant of parent glycoside hydrolase family 53 galactanase, useful
XX
XX PT in dairy industry, to prepare galacto-oligosaccharide and/or for
XX
XX PT hydrolysis of lactose.
XX
XX PS Example 2; Page; 262pp; English.
XX
XX CC The invention comprises variants of a glycoside hydrolase (GH) family 53
XX
XX CC galactanase. The GH family 53 galactanase variants of the invention are
XX
XX CC useful in the dairy industry for the preparation of galacto-
XX
XX CC oligosaccharide and the hydrolysis of lactose. The present amino acid
XX
XX CC sequence represents a mutant Corynascus heterothallicus GH family 53
XX
XX CC galactanase of the invention. NOTE: The present sequence is not shown in
XX
XX CC the specification, but has been created using the wild type Corynascus
XX
XX CC heterothallicus GH family 53 galactanase ADP84784 as a template.
XX
XX SQ Sequence 332 AA;
XX
XX Query Match 56.7%; Score 55; DB 8; Length 332;
XX
XX Best Local Similarity 53.3%; Pred. No. 1.7;
XX
XX Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGK 15
| : : : : :
Db 259 FPSDVKNIPFSPGQ 273

RESULT 5
AAW27063
ID AAW27063 standard; protein; 350 AA.
XX
XX AC AAW27063;
XX
XX DT 17-OCT-2003 (revised)
XX
XX DT 25-MAR-2003 (revised)
XX
XX DT 28-FEB-1998 (first entry)
XX
XX DE Myceliophthora thermophila galactanase.
XX
XX KW Galactanase; fruit juice; vegetable juice; wine; pectin; depectinisation;
XX
XX KW animal feed; foodstuff.
XX
XX OS Corynascus heterothallicus; strain CBS 117.65.
XX
XX FH Key Location/Qualifiers
XX
XX FT Peptide 1..18
XX
XX FT Peptide /label= Sig_peptide
XX
XX FT Protein 19..350
XX
XX FT Protein /label= Mat_protein
XX
XX FT Peptide 101..109
XX
XX FT Peptide /note= "Claim 7"
XX
XX FT Peptide 312..319
XX
XX FT Peptide /note= "Claim 7"
XX
XX
XX WO9732014-A1.
XX
XX PD 04-SEP-1997.
XX
XX

```

```

PF 28-FEB-1997; 97WO-DK0000092.
XX
XX PR 01-MAR-1996; 96DK-00000233.
XX
XX PR 01-MAR-1996; 96DK-00000235.
XX
XX PA (NOVO ) NOVO-NORDISK AS.
XX
XX PI Kofod LV, Kauppinen MS, Andersen LN, Clausen IG, Muellertz A;
XX
XX PI WPI; 1997-448686/41.
XX
XX DR N-PSDB; AAT85058.
XX
XX PT Fungal galactanase and related DNA - useful in animal feed industry and
XX
XX PT fruit juice depectinisation.
XX
XX PS Claim 23; Page 47-48; 67pp; English.
XX
XX CC This protein comprises Myceliophthora thermophila galactanase (EC-
XX
XX CC 3.2.1.89), as encoded by a cDNA clone (see AAT85058) insert in plasmid
XX
XX CC pYES 2.0, present in Saccharomyces cerevisiae DSM 9983. The galactanase
XX
XX CC catalyses the endohydrolysis of 1,4-beta-D- galactosidic linkages in
XX
XX CC arabinogalactans. It has a pH optimum of 6.0, a mol.wt. of 42 kDa, a pI
XX
XX CC of 7.8, a temperature optimum of 70 deg C, and a Km (% galactan) of 0.5-
XX
XX CC 0.9. The galactanase is useful in methods for the preparation of feed or
XX
XX CC food, for reducing the viscosity or water binding capacity of a plant
XX
XX CC wall-derived material, and in the production of wine or fruit or
XX
XX CC vegetable juice, especially the depectinisation of apple and pear juice.
XX
XX CC It can be expressed in transformed host cells and isolated in pure form,
XX
XX CC free of contaminating enzyme activities. 2 Motifs (see AAW27065-66) that
XX
XX CC are also found in Humicola insolens galactanase (see AAW27064) can be
XX
XX CC used to design primers (see AAT85060-61) useful in the isolation of
XX
XX CC galactanases from other fungi of the order Sordariales. (Updated on 25-
XX
XX CC MAR-2003 to correct PI field.) (Updated on 17-OCT-2003 to standardise OS
XX
XX CC field)
XX
XX SQ Sequence 350 AA;
XX
XX Query Match 56.7%; Score 55; DB 2; Length 350;
XX
XX Best Local Similarity 53.3%; Pred. No. 1.8;
XX
XX Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGK 15
| : : : : :
Db 277 FPSDVKNIPFSPGQ 291

RESULT 6
ADB70258
ID ADB70258 standard; protein; 846 AA.
XX
XX AC ADB70258;
XX
XX DT 04-DEC-2003 (first entry)
XX
XX DE C. neoformans amino acid sequence SEQ ID NO:3302.
XX
XX KW fungicide; gene therapy; infection.
XX
XX OS Cryptococcus neoformans.
XX
XX FN WO2003052076-A2.
XX
XX PD 26-JUN-2003.
XX
XX PF 17-DEC-2002; 2002WO-US040225.
XX
XX PR 17-DEC-2001; 2001US-0341261P.
XX
XX PA (ELIT-) ELITRA PHARM INC.
XX
XX PI Zamudio C, Eroshkin AM;
XX
XX DR WPI; 2003-533017/50.

```


XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX
PS Claim 1; SEQ ID NO 19046; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX

SQ Sequence 1136 AA;

Query Match 51.5%; Score 50; DB 8; Length 1136;
Best Local Similarity 64.3%; Pred. No. 50;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ERDISNVFSPDGK 15
| : : : : :
Db 648 ECTVNVVFPDGK 661

RESULT 9
ADS30836
ID ADS30836 standard; protein; 1155 AA.

XX ADS30836;

XX 02-DEC-2004 (first entry)

XX Bacterial polypeptide #19869.

XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

XX (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX

PS Claim 1; SEQ ID NO 19869; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX

SQ Sequence 1155 AA;

Query Match 49.5%; Score 48; DB 8; Length 1155;
Best Local Similarity 57.1%; Pred. No. 1.1e+02;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ERDISNVFSPDGK 15

Db 1108 QRGVNSVSPDGK 1121

RESULT 10

ADS23978

ID ADS23978 standard; protein; 699 AA.

XX ADS23978;

XX 02-DEC-2004 (first entry)

XX Bacterial polypeptide #13011.

XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.

OS Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.

XX PA (HINK/) HINKLE G J.

XX PA (SLAT/) SLATER S C.

XX PA (CHEN/) CHEN X.

XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX DR WPI; 2004-061375/06.

XX PT New recombinant DNA construct comprising a promoter positioned to provide

XX PT for expression of a polynucleotide encoding a polypeptide from a

XX PT microbial source, useful for producing plants with improved properties.

XX PS Claim 1; SEQ ID NO 13011; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a

XX CC promoter functional in a plant cell, where the promoter is positioned to

XX CC provide for expression of a polynucleotide encoding a polypeptide from a

XX CC microbial source. The invention also relates to a transformed plant

XX CC comprising the recombinant DNA construct and a method of producing a

XX CC transformed plant having an improved property. The plant is a crop plant

XX CC such as maize or soybean. The method of producing a transformed plant

XX CC having an improved property comprises transforming a plant with the

XX CC recombinant DNA construct and growing the transformed plant, where the

XX CC polynucleotide or polypeptide is useful for improving plant properties.

XX CC The recombinant DNA construct is useful for producing plants with

XX CC improved plant properties, e.g. improved cold, heat or drought tolerance,

XX CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

XX CC increased resistance to plant disease, better growth rate by modification

XX CC of the cell cycle pathway with plant growth regulators, increased rate of

XX CC homologous recombination, modified seed oil or protein yield and/or

XX CC content, improved yield by modification of carbohydrate, nitrogen or

XX CC phosphorus use and/or uptake, by modification of photosynthesis or by

XX CC providing improved plant growth and development under at least one stress

XX CC condition, improved lignin production or improved galactomannan

XX CC production. This sequence represents a bacterial polypeptide used in the

XX CC scope of the invention. Note: The sequence data for this patent did not

XX CC form part of the printed specification but was obtained in electronic

XX CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 699 AA;

Query Match 49.0%; Score 47.5; DB 8; Length 699;

Best Local Similarity 55.6%; Pred. No. 73;

Matches 10; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

QY 2 ERDISNVVF---SPDGK 16

Db 137 QRDIXNYPKVNKGDKP 154

RESULT 11

AAV75490

ID AAV75490 standard; protein; 134 AA.

XX AC AAV75490;

XX DT 21-MAR-2000 (first entry)

XX DE Neisseria meningitidis ORF 727 protein sequence SEQ ID NO:2454.

XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;

XX KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;

XX KW antibacterial; gene therapy.

XX OS Neisseria meningitidis.

XX PD

PN WO957280-A2.

XX PD 11-NOV-1999.

XX PF 30-APR-1999; 99WO-US009346.

XX PR 01-MAY-1998; 98US-0083758P.

XX PR 31-JUL-1998; 98US-0094869P.

XX PR 02-SEP-1998; 98US-0098994P.

XX PR 02-SEP-1998; 98US-0099062P.

XX PR 09-OCT-1998; 98US-0103749P.

XX PR 09-OCT-1998; 98US-0103794P.

XX PR 09-OCT-1998; 98US-0103796P.

XX PR 25-FEB-1999; 99US-0121528P.

XX PA (CHIR) CHIRON CORP.

XX PA (GENO-) INST GENOMIC RES.

XX PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;

XX PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;

XX PI Tettelin H, Venter JC;

XX DR WPI; 2000-062150/05.

XX DR N-PSDB; AAZ54252.

XX CC Novel Neisserial polypeptides predicted to be useful antigens for

XX CC vaccines and diagnostics.

XX PS Claim 2; Page 1173; 1453pp; English.

XX CC AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAZ74253 to AAY75941

XX CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides

XX CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent

XX CC PCR primers used in the exemplification of the present invention. The

XX CC polypeptides, the polynucleotides, antibodies and compositions of the

XX CC invention can be used as vaccines, as diagnostic reagents, and as

XX CC immunogenic compositions. The polypeptides can be used in the manufacture

XX CC of medicaments for treating or preventing infection due to Neisserial

XX CC bacteria (e.g. meningitis and septicaemia), to detect the presence of

XX CC Neisseria bacteria, or to raise antibodies. They may also be used to

XX CC screen for agonists or antagonists, which may themselves have use as

XX CC antibacterial agents. The polynucleotides of the invention may also be

XX CC used in gene therapy protocols

XX SQ Sequence 134 AA;

Query Match 48.5%; Score 47; DB 3; Length 134;

Best Local Similarity 50.0%; Pred. No. 12;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 ERDISNVFSPDGK 15

Db 99 ERDLCKIFFPDSR 112

RESULT 12

ABO65040

ID ABO65040 standard; protein; 256 AA.

XX AC ABO65040;

XX DT 29-JUL-2004 (first entry)

XX DE Klebsiella pneumoniae polypeptide seqid 11557.

XX KW Recombinant expression vector; transcription regulatory element;

XX KW Klebsiella pneumoniae protein; antibacterial; vaccine.

XX OS Klebsiella pneumoniae.

XX KW US6610836-B1.

XX PD 26-AUG-2003.

XX PS Claim 7; SEQ ID NO 106; 433pp; English.

XX CC This invention relates to novel isolated *Alloiococcus otitidis* (A.

XX CC *otitidis*) nucleic acid molecules and encoded proteins thereof.

XX CC Specifically, it refers to proteins that are essential for the growth and

XX CC survival of the gram-positive bacterium *A. otitidis*, and hence provide

XX CC novel antibacterial targets. The present invention describes

XX CC pharmaceutical compositions and antisense compounds that are useful for

XX CC inhibiting activity or expression of these proteins. Furthermore, it

XX CC provides diagnostics and therapeutics that can be used to ameliorate

XX CC diseases that are associated with *A. otitidis*, such as otitis media with

XX CC effusion (OME) and various bacterial infections. Accordingly, these

XX CC compositions exhibit both antibacterial and antiinfective activities.

XX CC This polypeptide is an *A. otitidis* protein sequence of the invention.

XX SQ Sequence 590 AA;

Query Match 48.5%; Score 47; DB 8; Length 590;

Best Local Similarity 63.6%; Pred. No. 73;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 7 NVFSPDGKPC 17

DB 63 NCPYSDDGEP 73

RESULT 15

ADS41817

ID ADS41817 standard; protein; 626 AA.

XX AC ADS41817;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #20247.

XX KW Recombinant DNA construct; transformed plant; improved plant property;

XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

XX KW pathogen tolerance; pest tolerance; plant disease resistance;

XX KW cell cycle pathway modification; plant growth regulator;

XX KW homologous recombination; seed oil yield; protein yield; carbohydrate;

XX KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

XX KW bacterial polypeptide.

XX OS Bacteria.

XX PN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.

XX PA (HINK/) HINKLE G J.

XX PA (SLAT/) SLATER S C.

XX PA (CHEN/) CHEN X.

XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX DR WPI; 2004-061375/06.

XX PT New recombinant DNA construct comprising a promoter positioned to provide

XX PT for expression of a polynucleotide encoding a polypeptide from a

XX PT microbial source, useful for producing plants with improved properties.

XX PS Claim 1; SEQ ID NO 20247; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a

XX CC promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a

CC microbial source. The invention also relates to a transformed plant

CC comprising the recombinant DNA construct and a method of producing a

CC such as maize or soybean. The method of producing a transformed plant

CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the

CC polynucleotide or polypeptide is useful for improving plant properties.

CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

CC increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of

CC homologous recombination, modified seed oil or protein yield and/or

CC content, improved yield by modification of carbohydrate, nitrogen or

CC phosphorus use and/or uptake, by modification of photosynthesis or by

CC providing improved plant growth and development under at least one stress

CC condition, improved lignin production or improved galactomannan

CC production. This sequence represents a bacterial polypeptide used in the

CC scope of the invention. Note: The sequence data for this patent did not

CC form part of the printed specification but was obtained in electronic

CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 626 AA;

Query Match 48.5%; Score 47; DB 8; Length 626;

Best Local Similarity 50.0%; Pred. No. 78;

Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

OY 2 ERDISNVFSPDGK 15

DB 389 EKEVNSVAFSPDGQ 402

Search completed: March 18, 2005, 19:10:14

Job time : 165.84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

```
Run on:      March 18, 2005, 18:59:50 ; Search time 41.48 Seconds
            (without alignments)
            30.594 Million cell updates
```

Title: US-10-773-274A-2

Perfect score: 97

Sequence: 1 FERDISNVFSPDGKPC 17

Scoring table: BLOSUM62

scoring table: `BDOSUM2`
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Fast-processing: Minimum Match 0% Maximum Match 100%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

```
Database :
issue_rating_mn:
1: /cgm2_6/ptodata/1/1aa/5A_COMB.pcp.*
2: /cgm2_6/ptodata/1/1aa/5B_COMB.pcp.*
3: /cgm2_6/ptodata/1/1aa/6A_COMB.pcp.*
4: /cgm2_6/ptodata/1/1aa/6B_COMB.pcp.*
5: /cgm2_6/ptodata/1/1aa/PCRTUS_COMB.pcp.*
6: /cgm2_6/ptodata/1/1aa/backfiles1.pcp.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	55	56.7	350	3	US-09-137-855-2	Sequence 2, Appli
2	55	56.7	350	4	US-09-723-5480-2	Sequence 2, Appli
3	47	48.5	256	4	US-09-489-039A-11557	Sequence 11557, A
4	47	48.5	642	3	US-08-911-393-4	Sequence 4, Appli
5	47	48.5	642	4	US-09-955-509-4	Sequence 4, Appli
6	46	47.4	149	4	US-09-902-540-14114	Sequence 14114, A
7	46	47.4	571	4	US-09-460-295B-13	Sequence 13, Appl
8	45	46.4	204	2	US-08-531-525-48	Sequence 48, Appl
9	45	46.4	204	2	US-08-718-270A-48	Sequence 48, Appl
10	45	46.4	275	4	US-09-538-092-958	Sequence 958, App
11	45	46.4	288	4	US-09-538-092-764	Sequence 764, App
12	44	45.4	548	4	US-09-252-991A-20793	Sequence 20793, A
13	44	45.4	552	4	US-09-460-295B-8	Sequence 8, Appli
14	43.5	44.8	429	4	US-09-489-039A-13801	Sequence 13801, A
15	43	44.3	184	4	US-09-949-016-7506	Sequence 7506, Ap
16	43	44.3	304	4	US-09-248-796A-14813	Sequence 14813, A
17	43	44.3	349	3	US-09-137-855-4	Sequence 4, Appli
18	43	44.3	349	4	US-09-723-548C-4	Sequence 4, Appli
19	43	44.3	606	4	US-09-460-295B-12	Sequence 12, Appl
20	43	44.3	637	4	US-09-902-540-10450	Sequence 10450, A
21	43	44.3	1783	4	US-09-362-336A-2	Sequence 2, Appli
22	43	44.3	1804	4	US-09-362-336A-4	Sequence 4, Appli
23	42	43.3	131	4	US-09-252-991A-27452	Sequence 27452, A
24	42	43.3	230	4	US-09-893-737-210	Sequence 210, App
25	42	43.3	426	4	US-09-902-540-14406	Sequence 14406, A
26	42	43.3	589	4	US-09-252-991A-28836	Sequence 28836, A
27	42	43.3	919	4	US-09-538-092-1336	Sequence 1336, Ap

ALIGNMENTS

RESULT 1

```

US-09-137-855-2
; Sequence 2, Application US/09137855B
; Patent No. 624237
; GENERAL INFORMATION:
; APPLICANT: Kofod, Lene V.
; APPLICANT: Kauppinen, Markus S.
; APPLICANT: Andersen, Lene N.
; APPLICANT: Clausen, Ib G.
; APPLICANT: Mullettz, Anette
; TITLE OF INVENTION: An Enzyme With Galacta
; FILE REFERENCE: 4686.204-US
; CURRENT APPLICATION NUMBER: US/09/137,855B
; CURRENT FILING DATE: 1998-08-21
; EARLIER APPLICATION NUMBER: 0233/96
; EARLIER FILING DATE: 1996-03-01
; EARLIER APPLICATION NUMBER: 0235/96
; EARLIER FILING DATE: 1996-03-01
; EARLIER APPLICATION NUMBER: PCT/DK97/00092
; EARLIER FILING DATE: 1997-02-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Myceliophthora thermophila
US-09-137-855-2

```

Query Match 56.7%; Score 55; DB 3; Length 350;

Qy 1 FERDISNVPFSPDGK 15
p'b 277 FPSDVKNIPESPEGO 291

RESULT 2

```

US-09-723-548C-2
; Sequence 2, Application US/09723548C
; Patent No. 6485954
; GENERAL INFORMATION:
; APPLICANT: Kofod, Lene
; APPLICANT: Kauppinen, Markus
; APPLICANT: Andersen, Lene
; APPLICANT: Clausen, Ib
; APPLICANT: Mulleretz, Anette
; TITLE OF INVENTION: An Enzyme with Galactanase Activity
; FILE REFERENCE: 4686-514-US
; CURRENT APPLICATION NUMBER: US/09/723,548C

```


<p>Db</p>	US-09-955-909-4 Query Match 48.5%; Score 47; DB 4; Length 642; Best Local Similarity 60.0%; Pred. No. 23; Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY	1 FERDISNVPFSPDGK 15 : : : :
Db	12 FRQSSKVFEKGDKG 26 : :
RESULT 6 US-09-902-540-14114 Sequence 14114, Application US/0902540 Patent No. 6833447 GENERAL INFORMATION: APPLICANT: Goldman, Barry S. APPLICANT: Hinkle, Gregory J. APPLICANT: Slater, Steven C. APPLICANT: Wiegand, Roger C. TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof FILE REFERENCE: 38-10(15849)B CURRENT APPLICATION NUMBER: US/09/902,540 CURRENT FILING DATE: 2001-07-10 PRIOR APPLICATION NUMBER: 60/217,883 PRIOR FILING DATE: 2000-07-10 NUMBER OF SEQ ID NOS: 16825 SEQ ID NO 14114 LENGTH: 149 TYPE: PRT ORGANISM: Myxococcus xanthus US-09-902-540-14114	
QY	1 FERDISNVPFSPDG 14 : : : :
Db	109 YEPDVEGVAFSPDG 122 : : : :
RESULT 7 US-09-460-295B-13 Sequence 13, Application US/09460295B Patent No. 6710030 GENERAL INFORMATION: APPLICANT: UNIVERSITY OF SOUTHERN CALIFORNIA TITLE OF INVENTION: CONTOPROSTATIN (CN) AND METHODS FOR ITS USE IN PREVENTING METASTASIS TITLE OF INVENTION: OTHER CONDITIONS FILE REFERENCE: 1279-338C3/09801388 CURRENT APPLICATION NUMBER: US/09/460,295B CURRENT FILING DATE: 1999-12-10 PRIOR APPLICATION NUMBER: US 09/163,047 PRIOR FILING DATE: 1998-09-29 NUMBER OF SEQ ID NOS: 15 SOFTWARE: PatentIn version 3.1 SEQ ID NO 13 LENGTH: 571 TYPE: PRT ORGANISM: Bothrops jararaca US-09-460-295B-13	
QY	1 FERDISNVPFSPDGK 15 : : : :

Db	38 FSKDYSEIHYSPDGR 52 : : : :
RESULT 8 US-08-531-525-48 Sequence 48, Application US/08531525 Patent No. 5840683 GENERAL INFORMATION: APPLICANT: Hlavka, Joseph J. APPLICANT: Pincus, Matthew R. APPLICANT: No. 5840683le, John F. APPLICANT: Abajian, Henry B. APPLICANT: Kende, Andrew S. TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action NUMBER OF SEQUENCES: 52 CORRESPONDENCE ADDRESS: ADDRESSEE: Greenlee and Winner, P.C. 201 STREET: 5370 Manhattan Circle, Suite 201 CITY: Boulder STATE: Colorado COUNTRY: US ZIP: 80303 COMPUTER READABLE FORM: MEDIUM TYPE: Floppy disk COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS SOFTWARE: PatentIn Release #1.0, Version #1.30 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/531,525 FILING DATE: 21-SEP-1995 CLASSIFICATION: 530 ATTORNEY/AGENT INFORMATION: NAME: Ferber, Donna M. REGISTRATION NUMBER: 33,878 REFERENCE/DOCKET NUMBER: 37-94 TELECOMMUNICATION INFORMATION: TELEPHONE: (303) 499-8080 TELEFAX: (303) 499-8089 INFORMATION FOR SEQ ID NO: 48: SEQUENCE CHARACTERISTICS: LENGTH: 204 amino acids TYPE: amino acid STRANDEDNESS: single TOPOLOGY: linear MOLECULE TYPE: protein HYPOTHETICAL: NO ORIGINAL SOURCE: ORGANISM: Caenorhabditis elegans US-08-531-525-48	
QY	9 PFPSPDGKPC 17
Db	193 PFKDDGKPC 201
RESULT 9 US-08-718-270A-48 Sequence 48, Application US/08718270A Patent No. 5910478 GENERAL INFORMATION: APPLICANT: Hlavka, Joseph J. APPLICANT: Pincus, Matthew R. APPLICANT: No. 5910478le, John F. APPLICANT: Abajian, Henry B. APPLICANT: Kende, Andrew S. TITLE OF INVENTION: Peptidomimetics Inhibiting the Oncogenic Action of P21 Ras	

```

;
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 204 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Caenorhabditis elegans
;
US-08-718-270A-48
Query Match 46.4%; Score 45; DB 2; Length 204;
Best Local Similarity 77.8%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 PFSPDGKPC 17
Db 193 PFKDDGKPC 201

RESULT 10
US-09-538-092-958
; Sequence 958, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 958
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

;
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 204 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Caenorhabditis elegans
;
US-08-718-270A-48
Query Match 46.4%; Score 45; DB 2; Length 204;
Best Local Similarity 77.8%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 PFSPDGKPC 17
Db 193 PFKDDGKPC 201

RESULT 10
US-09-538-092-958
; Sequence 958, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 958
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

;
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P19387
;
US-09-538-092-958
Query Match 46.4%; Score 45; DB 4; Length 275;
Best Local Similarity 46.7%; Pred. No. 18;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 ERDINSVPSPDGKPC 16
Db 212 DEDESQAPYDPNGKP 226

RESULT 11
US-09-538-092-764
; Sequence 764, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 764
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YOR362C
;
US-09-538-092-764
Query Match 46.4%; Score 45; DB 4; Length 288;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 DISNVPSPDGK 15
Db 9 DLSNSVSPDGK 20

RESULT 12
US-09-252-991A-20793
; Sequence 20793, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196-136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20793
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
;
US-09-252-991A-20793
Query Match 45.4%; Score 44; DB 4; Length 548;

```

```
Best Local Similarity 60.0%; Pred. No. 59;
Matches 6; Conservative 2; Mismatches 2; Indels 2; Gaps 0;

Qy 8 VPFSPDGKPC 17
Db 281 IPRQDGRPC 290

RESULT 13
US-09-460-295B-8
; Sequence 8, Application US/09460295B
; Patent No. 6710030
; GENERAL INFORMATION:
; APPLICANT: UNIVERSITY OF SOUTHERN CALIFORNIA
; TITLE OF INVENTION: CONTOKTOSTATIN (CN) AND METHODS FOR ITS USE IN PREVENTING METAST
; FILE REFERENCE: 1279-338C3/09801388
; CURRENT APPLICATION NUMBER: US/09/460,295B
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: US 09/163,047
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Trimeresurus gramineus
US-09-460-295B-8

Query Match 45.4%; Score 44; DB 4; Length 552;
Best Local Similarity 46.7%; Pred. No. 60;
Matches 7; Conservative 3; Mismatches 5; Indels 5; Gaps 0;

Qy 1 FERDISNVFSPDGK 15
Db 149 FSEDYSEIHYSFDR 163

RESULT 14
US-09-489-039A-13801
; Sequence 13801, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13801
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13801

Query Match 44.8%; Score 43.5; DB 4; Length 429;
Best Local Similarity 58.8%; Pred. No. 54;
Matches 10; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

Qy 2 ERDISNVFSPD-GKPC 17
Db 257 ERDIQTVAFPQGOHC 273

RESULT 15
US-09-949-016-7506
; Sequence 7506, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7506
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7506

Query Match 44.3%; Score 43; DB 4; Length 184;
Best Local Similarity 57.1%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 3 RDISNVFSPDGK 16
Db 159 RKIENVPTGPNKP 172

Search completed: March 18, 2005, 19:12:10
Job time : 42.48 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 19:02:16 ; Search time 125.12 Seconds
(without alignments)
44.910 Million cell updates/sec

Title: US-10-773-274A-2

Perfect score: 97

Sequence: 1 FERDISNVFSPDGKPC 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1401741 seqs, 330541175 residues

Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	1255	16	US-10-839-729-10
2	97	100.0	1255	16	US-10-839-729-11
3	97	100.0	1255	16	US-10-839-729-12
4	97	100.0	1255	16	US-10-839-729-13
5	97	100.0	1257	17	US-10-808-187-1044
6	51	52.6	846	15	US-10-320-797-3302
7	50	51.5	655	15	US-10-369-493-20066
8	50	51.5	1136	15	US-10-369-493-19046
9	48	49.5	1155	15	US-10-369-493-19869
10	47.5	49.0	629	15	US-10-369-493-13011
11	47	48.5	626	15	US-10-369-493-13011
12	47	48.5	642	9	US-09-955-909-4
13	47	48.5	642	16	US-10-798-625-4

14	46	47.4	120	14	US-10-083-357-1281	Sequence 1381, Ap
15	46	47.4	396	14	US-10-156-761-9368	Sequence 9368, Ap
16	46	47.4	406	14	US-10-156-761-13869	Sequence 13869, A
17	46	47.4	508	15	US-10-369-493-9131	Sequence 9131, Ap
18	46	47.4	571	14	US-10-439-532-13	Sequence 13, Appl
19	46	47.4	571	16	US-10-712-584-13	Sequence 13, Appl
20	46	47.4	613	15	US-10-369-493-11181	Sequence 11181, A
21	46	47.4	756	15	US-10-369-493-19945	Sequence 19945, A
22	45	46.4	207	15	US-10-369-493-5837	Sequence 5837, Ap
23	45	46.4	288	15	US-10-369-493-22433	Sequence 22433, A
24	45	46.4	296	15	US-10-425-114-67435	Sequence 67435, A
25	45	46.4	310	16	US-10-767-701-43515	Sequence 43515, A
26	45	46.4	377	16	US-10-602-898A-10	Sequence 10, Appl
27	45	46.4	540	16	US-10-437-963-103746	Sequence 103746, A
28	45	46.4	1118	15	US-10-369-493-18845	Sequence 18845, A
29	44.5	45.9	296	15	US-10-296-115-778	Sequence 778, App
30	44.5	45.9	412	16	US-10-322-281-681	Sequence 681, App
31	44.5	45.9	447	16	US-10-322-281-684	Sequence 684, App
32	44	45.4	27	14	US-10-057-789-157	Sequence 157, App
33	44	45.4	27	14	US-10-212-628-157	Sequence 157, App
34	44	45.4	395	15	US-10-369-493-6993	Sequence 6993, Ap
35	44	45.4	552	14	US-10-439-532-8	Sequence 8, Appli
36	44	45.4	552	16	US-10-712-584-8	Sequence 8, Appli
37	44	45.4	613	15	US-10-369-493-1544	Sequence 1544, Ap
38	44	45.4	613	15	US-10-369-493-1956	Sequence 1956, Ap
39	44	45.4	748	15	US-10-369-493-3245	Sequence 3245, Ap
40	43.5	44.8	323	15	US-10-424-599-246318	Sequence 246318, A
41	43	44.3	81	15	US-10-424-599-190647	Sequence 190647, A
42	43	44.3	116	15	US-10-424-599-274941	Sequence 274941, A
43	43	44.3	147	16	US-10-767-701-36157	Sequence 36157, A
44	43	44.3	167	16	US-10-767-701-40448	Sequence 40448, A
45	43	44.3	183	9	US-09-925-300-1279	Sequence 1279, Ap

ALIGNMENTS

RESULT 1
US-10-839-729-10
; Sequence 10, Application US/10839729
; Publication No. US20050002953A1
; GENERAL INFORMATION:
; APPLICANT: Jens Herold
; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES
; TITLE OF INVENTION: AND METHODS OF USE
; FILE REFERENCE: BIOBANK.013A
; CURRENT APPLICATION NUMBER: US/10/839, 729
; PRIOR FILING DATE: 2004-05-04
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: SARS Coronavirus
US-10-839-729-10

Query Match 100.0%; Score 97; DB 16; Length 1255;
Best Local Similarity 100.0%; Pred. No. 5.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGKPC 17
Db 451 FERDISNVFSPDGKPC 467

RESULT 2
US-10-839-729-11
; Sequence 11, Application US/10839729
; Publication No. US20050002953A1
; GENERAL INFORMATION:
; APPLICANT: Jens Herold

```

; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES
; TITLE OF INVENTION: AND METHODS OF USE
; FILE REFERENCE: BIOBANK.013A
; CURRENT APPLICATION NUMBER: US/10/839,729
; PRIOR FILING DATE: 2004-05-04
; PRIOR APPLICATION NUMBER: 60/468703
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1255
; TYPE: PR1
; ORGANISM: SARS Coronavirus
US-10-839-729-11

```

```

Query Match 100.0%; Score 97; DB 16; Length 1255;
Best Local Similarity 100.0%; Pred. No. 5.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 FERDISNVFSPDGKPC 17
Db 451 FERDISNVFSPDGKPC 467

```

```

RESULT 3
US-10-839-729-12
; Sequence 12, Application US/10839729
; Publication No. US20050002953A1
; GENERAL INFORMATION:
; APPLICANT: Jens Herold
; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES
; FILE REFERENCE: BIOBANK.013A
; CURRENT APPLICATION NUMBER: US/10/839,729
; PRIOR FILING DATE: 2004-05-04
; PRIOR APPLICATION NUMBER: 60/468703
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1255
; TYPE: PR1
; ORGANISM: SARS Coronavirus
US-10-839-729-12

```

```

Query Match 100.0%; Score 97; DB 16; Length 1255;
Best Local Similarity 100.0%; Pred. No. 5.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 FERDISNVFSPDGKPC 17
Db 451 FERDISNVFSPDGKPC 467

```

```

RESULT 4
US-10-839-729-13
; Sequence 13, Application US/10839729
; Publication No. US20050002953A1
; GENERAL INFORMATION:
; APPLICANT: Jens Herold
; TITLE OF INVENTION: SARS-CORONAVIRUS VIRUS-LIKE PARTICLES
; FILE REFERENCE: BIOBANK.013A
; CURRENT APPLICATION NUMBER: US/10/839,729
; PRIOR FILING DATE: 2004-05-04
; PRIOR APPLICATION NUMBER: 60/468703
; PRIOR FILING DATE: 2003-05-06
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1255
; TYPE: PR1
; ORGANISM: SARS Coronavirus

```

US-10-839-729-13

```

Query Match 100.0%; Score 97; DB 16; Length 1255;
Best Local Similarity 100.0%; Pred. No. 5.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 FERDISNVFSPDGKPC 17
Db 451 FERDISNVFSPDGKPC 467

```

RESULT 5

```

US-10-808-187-1044
; Sequence 1044, Application US/10808187
; Publication No. US2005000909A1
; GENERAL INFORMATION:
; APPLICANT: PEIRIS, JOSEPH S. M.
; APPLICANT: YUEN, KWOK YUNG
; APPLICANT: POON, LIT MAN
; APPLICANT: GUAN, YI
; APPLICANT: CHAN, KWOK HUNG
; APPLICANT: NICHOLLS, JOHN
; TITLE OF INVENTION: A DIAGNOSTIC ASSAY FOR THE HUMAN VIRUS CAUSING SEVERE ACUTE
; FILE REFERENCE: V9661.0078
; CURRENT APPLICATION NUMBER: US/10/808,187
; CURRENT FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: 60/457,031
; PRIOR FILING DATE: 2003-03-24
; PRIOR APPLICATION NUMBER: 60/457,730
; PRIOR FILING DATE: 2003-03-26
; PRIOR APPLICATION NUMBER: 60/459,931
; PRIOR FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: 60/460,357
; PRIOR FILING DATE: 2003-04-03
; PRIOR APPLICATION NUMBER: 60/461,265
; PRIOR FILING DATE: 2003-04-08
; PRIOR APPLICATION NUMBER: 60/462,805
; PRIOR FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 60/468,139
; PRIOR FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: 60/464,886
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/471,200
; PRIOR FILING DATE: 2003-05-16
; NUMBER OF SEQ ID NOS: 2476
; SOFTWARE: Patent in ver. 3.2
; SEQ ID NO 1044
; LENGTH: 1257
; TYPE: PR1
; ORGANISM: Human severe acute respiratory system virus
US-10-808-187-1044

```

```

Query Match 100.0%; Score 97; DB 17; Length 1257;
Best Local Similarity 100.0%; Pred. No. 5.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 FERDISNVFSPDGKPC 17
Db 453 FERDISNVFSPDGKPC 469

```

RESULT 6

```

US-10-320-797-3302
; Sequence 3302, Application US/10320797
; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Eroshkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797

```

; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3302
; LENGTH: 846
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (556)..(556)
; OTHER INFORMATION: xaa = any amino acid
US-10-320-797-3302

Query Match 52.6%; Score 51; DB 15; Length 846;
Best Local Similarity 60.0%; Pred. No. 40;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGK 15
| : : : : :
Db 95 FKRKVHVSVFSPDGK 109

RESULT 7

US-10-369-493-20066
; Sequence 20066, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 20066
; LENGTH: 655
; TYPE: PRT
; ORGANISM: No. US20030233675A1toc punctiforme
US-10-369-493-20066

Query Match 51.5%; Score 50; DB 15; Length 655;
Best Local Similarity 53.3%; Pred. No. 43;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGK 15
| : : : : :
Db 600 YEELTSVSVFSPDGK 614

RESULT 8

US-10-369-493-19046
; Sequence 19046, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19046
; LENGTH: 1136
; TYPE: PRT
; ORGANISM: Anabaena PCC7120
US-10-369-493-19046

Query Match 51.5%; Score 50; DB 15; Length 1136;
Best Local Similarity 64.3%; Pred. No. 78;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ERDISNVFSPDGK 15
| : : : : :
Db 648 EQTVNNVVFSPDGK 661

RESULT 9

US-10-369-493-19869
; Sequence 19869, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19869
; LENGTH: 1155
; TYPE: PRT
; ORGANISM: No. US20030233675A1toc punctiforme
US-10-369-493-19869

Query Match 49.5%; Score 48; DB 15; Length 1155;
Best Local Similarity 57.1%; Pred. No. 1.6e+02;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ERDISNVFSPDGK 15
| : : : : :
Db 1108 QGVNSVSVFSPDGK 1121

RESULT 10

US-10-369-493-13011
; Sequence 13011, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 13011
; LENGTH: 699
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
; FEATURE:

NAME/KEY: unsure
LOCATION: (1)..(699)
OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-13011

Query Match 49.0%; Score 47.5; DB 15; Length 699;
Best Local Similarity 55.6%; Pred. No. 1.1e+02;
Matches 10; Conservative 2; Mismatches 3; Indels 3; Gaps 1;

Qy 2 ERDISNVPF---SPDKP 16
Db 137 QRDINKVPFKVKNKDKP 154

RESULT 11
US-10-369-493-20247
Sequence 20247, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xiaofeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 20247
LENGTH: 626
TYPE: PRT
ORGANISM: No. US20030233675A1toc punctiforme
US-10-369-493-20247

Query Match 48.5%; Score 47; DB 15; Length 626;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Qy 2 ERDISNVPSPDGK 15
Db 389 EKEVNSVAFSPDGQ 402

RESULT 12
US-09-955-909-4
Sequence 4, Application US/09955909
Patent No. US20020150995A1
GENERAL INFORMATION:
APPLICANT: PELLETIER, Marc
BARKER, William A.
HAKES, David J.
ZOPE, David A.
TITLE OF INVENTION: METHODS FOR PRODUCING SIALYLLOLIGOSACCHARIDES IN A DAIRY SOURCE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/798,625
FILING DATE: 11-Mar-2004
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/955,909
FILING DATE: 18-Sep-2001
APPLICATION NUMBER: 08/911,393
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7188-032-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)7909090
TELEFAX: (212)8699741
TELEX: 66141 PENNIE

FILING DATE: 18-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/911,393
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7188-032-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)7909090
TELEFAX: (212)8699741
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 642 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-955-909-4

Query Match 48.5%; Score 47; DB 9; Length 642;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FERDISNVPSPDGK 15
Db 12 FKQSSKVPFEKDGK 26

RESULT 13
US-10-798-625-4
Sequence 4, Application US/10798625
Publication No. US20040185146A1
GENERAL INFORMATION:
APPLICANT: PELLETIER, Marc
BARKER, William A.
HAKES, David J.
ZOPE, David A.
TITLE OF INVENTION: METHODS FOR PRODUCING SIALYLLOLIGOSACCHARIDES IN A DAIRY SOURCE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/798,625
FILING DATE: 11-Mar-2004
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/955,909
FILING DATE: 18-Sep-2001
APPLICATION NUMBER: 08/911,393
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7188-032-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)7909090
TELEFAX: (212)8699741
TELEX: 66141 PENNIE

```

; INFORMATION FOR SEQ ID NO: 4:
;   SEQUENCE CHARACTERISTICS:
;       LENGTH: 642 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-798-625-4

```

```
Query Match      48.5%; Score 47; DB 16; Length 642;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```

Qy      1 FERDISNVFFSPDGK 15
         | : | | | | | |
Db     12 FKROSSKVPFEKDGK 26

```

```

RESULT 14
US-10-083-357-1281
; Sequence 1281, Application US/10083357
; Publication No. US20030054370A1
; GENERAL INFORMATION:
; APPLICANT: Qiangdong Zeng et al.
; TITLE OF INVENTION: Systemic Discovery of New Genes
; FILE REFERENCE: 032796-090
; CURRENT APPLICATION NUMBER: US/10/083,357
; CURRENT FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 1346
; SEQ ID NO 1281
; LENGTH: 120
; TYPE: Prt
; ORGANISM: Saccharomyces cerevisiae
US-10-083-357-1281

```

Query Match 47.4%; Score 46; DB 14; Length 120;
Best Local Similarity 64.3%; Pred. No. 29;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 ERDISNVFSPDGK 15
 : | | | | | | |
D6 8 QRKISAVKFSPDGR 21

```

RESULT 15
US-10-156-761-9368
; Sequence 9368, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9368
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9368

```

Query Match 47.4%; Score 46; DB 14; Length 396;

```

Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 4 DISNVFSPDGKPC 17
    |||::|||
    ||::|||
Db 236 DIGHIOAVDGRPC 249

```

Search completed: March 18, 2005, 19:15:21
Job time : 126.12 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 18:59:20 ; Search time 33.32 Seconds
(without alignments)
49.090 Million cell updates/sec

Title: US-10-773-274A-2
Perfect score: 97
Sequence: 1 FERDISNVFSPDGKPC 17
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	51.5	1683	2 AF2071	WD-40 repeat prote
2	48	49.5	736	2 D96830	probable heat-shoc
3	48	49.5	831	2 E36830	hypothetical prote
4	47.5	49.0	570	2 C90485	beta-glucuronidase
5	47	48.5	134	2 E81133	hypothetical prote
6	47	48.5	676	2 AH2195	hypothetical prote
7	46	47.4	132	1 MFIV2J	matrix protein M2
8	46	47.4	571	2 S24789	jaraahagin C precu
9	46	47.4	831	2 S74252	heat shock protein
10	46	47.4	1329	2 AE1901	WD-repeat containi
11	45.5	46.9	377	2 AF2273	hypothetical prote
12	45	46.4	207	2 S40747	hypothetical prote
13	45	46.4	275	2 A36264	DNA-directed RNA p
14	45	46.4	275	2 JG6181	RNA polymerase II
15	45	46.4	288	1 SBNYCI	proteasome endopep
16	45	46.4	337	2 D92884	conserved hypothet
17	45	46.4	934	2 AG1889	WD-40 repeat prote
18	45	46.4	961	2 E86245	hypothetical prote
19	44	45.4	395	2 T23317	hypothetical prote
20	44	45.4	480	1 A30065	trigraamin precuro
21	44	45.4	586	2 T38992	WD-40 repeat regul
22	44	45.4	594	2 S50802	hypothetical prote
23	44	45.4	613	2 S50721	dnaK-type molecula
24	44	45.4	613	2 S20149	dnaK-type molecula
25	43.5	44.8	470	2 C37032	Fe-S oxidoreductas
26	43	44.3	58	2 S49347	ccoQ protein - Rho
27	43	44.3	103	2 G84741	hypothetical prote
28	43	44.3	355	2 T30071	hypothetical prote
29	43	44.3	407	2 S66260	metalloproteinase

atrolysin C (EC 3.
atrolysin C (EC 3.
atrolysin B (EC 3.
serine-type D-Ala-
fibrinolytic metal
metalloproteinase-
catrocollastatin p
hypothetical prote
WD-40 repeat prote
WD-40 repeat prote
probable GTPase-ac
exopolyphosphatase
hypothetical prote
guanylate kinase T
WD-40 repeat prote
glyceraldehyde-3-p

ALIGNMENTS

RESULT 1

AF2071 WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C;Accession: AF2071

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi

Nakazaki, N.; Shimura, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AF2071

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1683 <KUR>

A;Cross-references: UNIPROT:Q8YV57; GB:BA000019; PIDN:BAB73823.1; PID:gi17131215; GSPDB:G

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: ali2124

Query Match 51.5%; Score: 50; DB 2; Length 1683;
Best Local Similarity 64.3%; Pred. No. 15;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ERDISNVFSPDGK 15

Db 1155 EQTVNNYFSPDGK 1168

RESULT 2

D96830 probable heat-shock protein, 41956-44878 [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C;Accession: D96830

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federapfel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Dewar, K.,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis .

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: D96830

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-736 <STO>

A;Cross-references: UNIPROT:Q9CA95; GB:AE005173; NID:G6453874; PIDN:AAF09058.1; GSPDB:GN

C:Genetics:

A:Gene: F19K16.12

A:Map position: 1

C:Superfamily: heat shock protein 91

Query Match 49.5%; Score 48; DB 2; Length 736;

Best Local Similarity 52.6%; Pred. No. 13;

Matches 10; Conservative 3; Mismatches 2; Indels 4; Gaps 1;

QY 2 ERDISNVFPS----PDGKP 16

:||||:|||||

Db 81 QRDIKSLPFSVTEGPDGYP 99

RESULT 3

E96830

hypothetical protein F18B13.1 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: E96830

R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E96830

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-831 <STO>

A:Cross-references: UNIPROT:Q9S7C0; GB:AE005173; NID:g5902359; PIDN:AA55461.1; GSPDB:GN

C:Genetics:

A:Gene: F18B13.1

A:Map position: 1

C:Superfamily: heat shock protein 91

Query Match 49.5%; Score 48; DB 2; Length 831;

Best Local Similarity 52.6%; Pred. No. 15;

Matches 10; Conservative 3; Mismatches 2; Indels 4; Gaps 1;

QY 2 ERDISNVFPS----PDGKP 16

:||||:|||||

Db 81 QRDIKSLPFSVTEGPDGYP 99

RESULT 4

C90485

beta-glucuronidase (gusB) [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C:Accession: C90485

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Ayez, M.J.; Chan-

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, H.

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: C90485

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-570 <KUR>

A:Cross-references: UNIPROT:Q97U11; GB:AE006641; NID:gl3816434; PIDN:AAK43138.1; GSPDB:G

C:Genetics:

A:Gene: gusB

C:Superfamily: beta-glucuronidase

Query Match 49.0%; Score 47.5; DB 2; Length 570;

Best Local Similarity 64.3%; Pred. No. 12;

C:Genetics:

A:Gene: F19K16.12

A:Map position: 1

C:Superfamily: heat shock protein 91

Query Match 49.5%; Score 48; DB 2; Length 736;

Best Local Similarity 52.6%; Pred. No. 13;

Matches 10; Conservative 3; Mismatches 2; Indels 4; Gaps 1;

QY 2 ERDISNVFPS----PDGKP 16

:||||:|||||

Db 81 QRDIKSLPFSVTEGPDGYP 99

RESULT 5

E81133

hypothetical protein NMB0990 [imported] - Neisseria meningitidis (strain MC58 serogroup

C:Species: Neisseria meningitidis

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004

C:Accession: E81133

R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve;

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755; PMID:10710307

A:Accession: E81133

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-134 <TET>

A:Cross-references: UNIPROT:Q9U2M1; GB:AE002450; GB:AE002098; NID:g7226229; PIDN:AAF4139

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB0990

C:Superfamily: Neisseria meningitidis hypothetical protein NMB1191

Query Match 48.5%; Score 47; DB 2; Length 134;

Best Local Similarity 50.0%; Pred. No. 2.8;

Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

RESULT 6

AH2195

hypothetical protein alr3119 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C:Accession: AH2195

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.;

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AH2195

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-676 <KUR>

A:Cross-references: UNIPROT:Q8YSG6; GB:BA000019; PIDN:BAB74818.1; PID:gl7132214; GSPDB:G

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr3119

Query Match 48.5%; Score 47; DB 2; Length 676;

Best Local Similarity 72.7%; Pred. No. 17;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 7

MFIV2J

matrix protein M2 - influenza C virus (strain C/JJ/50) (fragment)

C:Species: influenza C virus

C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 21-Nov-1997
C;Accession: B28878
R;Yamashita, M.; Krystal, M.; Palese, P.
J. Virol. 62, 3348-3355, 1988
A;Title: Evidence that the matrix protein of influenza C virus is coded for by a spliced
A;Reference number: A93036; MUID:88300888; PMID:3404579
A;Accession: B28878
A;Molecule type: mRNA
A;Residues: 1-132 <YAM>
A;Cross-references: GB:M22038; NID:G325303
A;Note: this ORF is not annotated in GenBank entry FLCCJMJ
C;Genetics:
A;Map position: segment 6
C;Superfamily: influenza C virus matrix protein M2
C;Keywords: matrix protein

Query Match 47.4%; Score 46; DB 1; Length 132;
Best Local Similarity 69.2%; Pred. No. 4;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ERDISNVFSPDG 14
Db 116 ETDISPFPFSDNG 128
|||||:|||||

RESULT 8
S24789
Jararhagin C precursor - jararaca (fragment)
N;Alternate names: single chain botrocetin
N;Contains: disintegrin-like 28K protein; hemorrhagic proteinase (EC 3.4.24.-)
C;Species: Bothrops jararaca (jararaca)
C;Date: 20-Feb-1995 #sequence_revision 29-Aug-1997 #text_change 09-Jul-2004
C;Accession: S24789; JC2245; A44463; A37958; JC2373
R;Paine, M.J.I.
Submitted to the EMBL Data Library, August 1992
A;Reference number: S24789
A;Accession: S24789
A;Molecule type: mRNA
A;Residues: 1-571 <PAI>
A;Cross-references: UNIPROT:P30431; EMBL:X68251; NID:G62467; PID:G62468
R;Usami, Y.; Fujimura, Y.; Miura, S.; Shima, H.; Yoshida, E.; Yoshioka, A.; Hirano, K.;
Biochem. Biophys. Res. Commun. 201, 331-339, 1994
A;Title: A 28 kDa-protein with disintegrin-like structure (jararhagin-C) purified from B
A;Reference number: JC2245; MUID:94256999; PMID:8198592
A;Accession: JC2245
A;Molecule type: Protein
A;Residues: 360-571 <USA>
A;Experimental source: venom
R;Paine, M.J.; Desmond, H.P.; Theakston, R.D.; Crampton, J.M.
J. Biol. Chem. 267, 22869-22876, 1992
A;Title: Purification, cloning, and molecular characterization of a high molecular weight
ily.
A;Reference number: A44463; MUID:93054601; PMID:1385408
A;Accession: A4463
A;Molecule type: mRNA
A;Residues: 1-23,'Q',25-92,'G',94-131,'G',133-169,'Q',171-571 <PA2>
A;Cross-references: GB:X68251
A;Experimental source: venom gland
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIP:118104)
R;Fujimura, Y.; Titani, K.; Usami, Y.; Suzuki, M.; Oyama, R.; Matsui, T.; Fukui, H.; Sug
Biochemistry 30, 1957-1964, 1991
A;Title: Isolation and chemical characterization of two structurally and functionally di
A;Reference number: A37958; MUID:91129280; PMID:1993206
A;Accession: A37958
A;Molecule type: protein
A;Residues: 360-372,'E',374-378,'X',380-384 <FUJ>
A;Note: 361-Val was also found
C;Comment: Inhibits collagen- and ADP-induced platelet aggregation.
C;Superfamily: mouse meltrin alpha; disintegrin homology
C;Keywords: hydrolase; metalloproteinase; venom; zinc
F;360-571/Product: jararhagin C #status experimental <MAT>
F;362-444/Domain: disintegrin homology <DIS>

F;295,299,305/Binding site: zinc (His) #status predicted
F;296/Active site: Glu #status predicted

Query Match 47.4%; Score 46; DB 2; Length 571;
Best Local Similarity 46.7%; Pred. No. 20;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGK 15
Db 38 FSKDYSEIHYSPOGR 52
|||||:|||||

RESULT 9
S74252
heat shock protein 91 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 28-Oct-1996 #sequence_revision 14-Nov-1997 #text_change 09-Jul-2004
C;Accession: S74252
R;Storozhenko, S.; de Pauw, P.; Kushnir, S.; van Montagu, M.; Inze, D.
FEBS Lett. 390, 113-118, 1996
A;Title: Identification of an Arabidopsis thaliana cDNA encoding a HSP70-related protein;
A;Reference number: S74252; MUID:96314561; PMID:8706819
A;Accession: S74252
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-831 <STO>
A;Cross-references: UNIPROT:Q96269; EMBL:Z70314; NID:G1495250; PID:CAA94389.1; PID:G149
C;Genetics:
A;Gene: HSP91
C;Superfamily: heat shock protein 91
C;Keywords: heat shock

Query Match 47.4%; Score 46; DB 2; Length 831;
Best Local Similarity 52.6%; Pred. No. 31;
Matches 10; Conservative 2; Mismatches 3; Indels 4; Gaps 1;

Qy 2 ERDISNVFSS----PDGKP 16
Db 81 QRDIKSFFFSVTEGPDGVP 99
|||||:|||||

RESULT 10
AE1901
WD-repeat containing protein [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AE1901
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AE1901; MUID:21595285; PMID:11759840
A;Accession: AE1901
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1329 <KUR>
A;Cross-references: UNIPROT:Q8YVY7; GB:BA000019; PIDN:BA072716.1; PID:G17130104; GSPDB:C
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all0759

Query Match 47.4%; Score 46; DB 2; Length 1329;
Best Local Similarity 53.8%; Pred. No. 53;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RDISNVFSPDGK 15
Db 649 KEVNDVAFSPDGK 661
|||||:|||||

RESULT 11
AF2273

hypothetical protein all3741 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AF2273
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Rep. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF2273
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-377 <KUR>
A:CROSS-references: UNIPROT:Q8YQ88; GB:BA000019; PIDN:BA075440.1; PID:gl17132875; GSPDB:GN0001
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all3741
C:Superfamily: Campylobacter jejuni hypothetical protein Cj1270c

Query Match 46.9%; Score 45.5; DB 2; Length 377;
Best Local Similarity 66.7%; Pred. No. 15;

Matches 10; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 3 RDISNVFPSPDGKPC 17

Db 276 KDIVTP-SPVGKPC 289

RESULT 12

S40747
hypothetical protein F54C8.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 09-Jul-2004
C:Accession: S40747
R:Berkas, M.; Smith, A.
submitted to the EMBL Data Library, February 1992
A:Reference number: S40473
A:Accession: S40747
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-207 <BER>
A:CROSS-references: UNIPROT:P34443; EMBL:Z22178; NID:g297949; PID:g297954
C:Genetics:
A:Introns: 25/1; 67/1; 99/2; 118/2; 160/3
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; nucleotide binding; P-loop
F:20-27/Region: nucleotide-binding motif A (P-loop)
F:126-129/Region: GTP-binding NKXD motif

Query Match 46.4%; Score 45; DB 2; Length 207;
Best Local Similarity 77.8%; Pred. No. 9.5;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 9 PFSPDGKPC 17

Db 196 PFKDDGKPC 204

RESULT 13

A36264
DNA-directed RNA polymerase (EC 2.7.7.6) II 33K chain - human
C:Species: Homo sapiens (man)
C>Date: 18-Jan-1991 #sequence_revision 18-Jan-1991 #text_change 09-Jul-2004
C:Accession: A36264
R:Patl, U.K.; Weissman, S.M.
J. Biol. Chem. 265, 8400-8403, 1990
A:Title: The amino acid sequence of the human RNA polymerase II 33-kDa subunit hRPB 33
A:Reference number: A36264; MUID:90256750; PMID:2187864
A:Accession: A36264
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-275 <PAT>

A:CROSS-references: UNIPROT:P19387; GB:J05448; NID:g337496; PIDN:AAA36586.1; PID:g337497
C:Superfamily: DNA-directed RNA polymerase II, RPB3 subunit; ferredoxin 2 [4Fe-4S] homolom
C:Keywords: nucleotidyltransferase; transcription

Query Match 46.4%; Score 45; DB 2; Length 275;
Best Local Similarity 46.7%; Pred. No. 13;

Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 ERDISNVFPSPDGKPC 16

Db 212 DEDESQAPYDPNGKP 226

RESULT 14

JC6181
RNA polymerase II 31.3K chain - mouse
C:Species: Mus musculus (house mouse)
C>Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C:Accession: JC6181
R:Korobko, I.V.; Yamamoto, K.; Nogi, Y.; Muramatsu, M.
Gene 185, 1-4, 1997
A:Title: Protein interaction cloning in yeast of the mouse third largest RNA polymerase
A:Reference number: JC6181; MUID:97186689; PMID:9034305
A:Accession: JC6181
A:Molecule type: mRNA
A:Residues: 1-275 <KOR>
A:CROSS-references: UNIPROT:P97760; DDBJ:D83999; NID:gl850788; PIDN:BA012205.1; PID:gl850789
C:Superfamily: DNA-directed RNA polymerase II, RPB3 subunit; ferredoxin 2 [4Fe-4S] homolom

Query Match 46.4%; Score 45; DB 2; Length 275;
Best Local Similarity 46.7%; Pred. No. 13;

Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 ERDISNVFPSPDGKPC 16

Db 212 DEDESQAPYDPNGKP 226

RESULT 15

SNBYC1
proteasome endopeptidase complex (EC 3.4.25.1) chain YC1 - yeast (Saccharomyces cerevisi
N:Alternate names: proteasome chain YC1; protein O6650; protein YOR362C
C:Species: Saccharomyces cerevisiae
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C:Accession: S11182; A38769; S67274
R:Fujiwara, T.; Tanaka, K.; Orino, E.; Yoshimura, T.; Kumatori, A.; Tamura, T.; Chung, C.
J. Biol. Chem. 265, 16604-16613, 1990
A:Title: Proteasomes are essential for yeast proliferation. cDNA cloning and gene disrupt
A:Reference number: S11182; MUID:90375535; PMID:1697860
A:Accession: S11182
A:Molecule type: mRNA
A:Residues: 1-288 <FUJ>
A:CROSS-references: UNIPROT:P21242; EMBL:M55436; NID:gl73203; PIDN:AAA35227.1; PID:gl73203

Query Match 46.4%; Score 45; DB 2; Length 275;
Best Local Similarity 46.7%; Pred. No. 13;

Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 ERDISNVFPSPDGKPC 16

Db 212 DEDESQAPYDPNGKP 226

RESULT 15

SNBYC1
proteasome endopeptidase complex (EC 3.4.25.1) chain YC1 - yeast (Saccharomyces cerevisi
N:Alternate names: proteasome chain YC1; protein O6650; protein YOR362C
C:Species: Saccharomyces cerevisiae
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004
C:Accession: S11182; A38769; S67274
R:Fujiwara, T.; Tanaka, K.; Orino, E.; Yoshimura, T.; Kumatori, A.; Tamura, T.; Chung, C.
J. Biol. Chem. 265, 16604-16613, 1990
A:Title: Proteasomes are essential for yeast proliferation. cDNA cloning and gene disrupt
A:Reference number: S11182; MUID:90375535; PMID:1697860
A:Accession: S11182
A:Molecule type: mRNA
A:Residues: 1-288 <FUJ>
A:CROSS-references: UNIPROT:P21242; EMBL:M55436; NID:gl73203; PIDN:AAA35227.1; PID:gl73203

Query Match 46.4%; Score 45; DB 1; Length 288;
Best Local Similarity 66.7%; Pred. No. 14;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 DISNVFSPDGK 15
|:|:|:|:|:|:
Db 9 DLSNVFSPDGR 20

Search completed: March 18, 2005, 19:11:07
Job time : 35.32 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 18:50:09 ; Search time 155.72 Seconds
(without alignments)
55.904 Million cell updates/sec

Title: US-10-773-274A-2

Perfect score: 97

Sequence: 1 FERDISNVFPSPDGKPC 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	1255	1	VGL2_CVHSA
2	97	100.0	1255	2	P59594 human coron
3	97	100.0	1255	2	Q692E4
4	97	100.0	1255	2	Q6DEU4 sars corona
5	97	100.0	1255	2	Q6GYR1 sars corona
6	97	100.0	1255	2	Q6JH38 sars corona
7	97	100.0	1255	2	Q6QU82 sars corona
8	97	100.0	1255	2	Q6R7Y6 sars corona
9	97	100.0	1255	2	Q6RCW5 sars corona
10	97	100.0	1255	2	Q6RCX6 sars corona
11	97	100.0	1255	2	Q6RCY7 sars corona
12	97	100.0	1255	2	Q6RCZ8 sars corona
13	97	100.0	1255	2	Q6RD09 sars corona
14	97	100.0	1255	2	Q6RD20 sars corona
15	97	100.0	1255	2	Q6RD31 sars corona
16	97	100.0	1255	2	Q6RD42 sars corona
17	97	100.0	1255	2	Q6RD53 sars corona
18	97	100.0	1255	2	Q6RD64 sars corona
19	97	100.0	1255	2	Q6S8D7 sars corona
20	97	100.0	1255	2	Q6T7X4 sars corona
21	97	100.0	1255	2	Q6T7X5 sars corona
22	97	100.0	1255	2	Q6T7X6 sars corona
23	97	100.0	1255	2	Q6T7X7 sars corona
24	97	100.0	1255	2	Q6T7X8 sars corona
25	97	100.0	1255	2	Q6T7X9 sars corona
26	97	100.0	1255	2	Q6T7Y0 sars corona
27	97	100.0	1255	2	Q6T7Y1 sars corona
28	97	100.0	1255	2	Q6T7P8 sars corona
29	97	100.0	1255	2	Q6UZF0 sars corona
30	97	100.0	1255	2	Q6UZF4 sars corona
31	97	100.0	1255	2	Q6V585 sars corona

32	97	100.0	1255	2	Q6VA78
33	97	100.0	1255	2	Q6VA89 sars corona
34	97	100.0	1255	2	Q6VAA0 sars corona
35	97	100.0	1255	2	Q6WGP3 sars corona
36	55	56.7	332	2	P83692 thielavia h
37	52	53.6	1233	2	Q8TMM4
38	50	51.5	310	2	Q6FY54
39	50	51.5	1683	1	YL24_ANASP
40	49	50.5	1029	2	Q6FPT5
41	49	50.5	2322	2	Q6UDW6
42	48	49.5	736	2	Q9CA95
43	48	49.5	831	2	Q8VZ83
44	48	49.5	831	2	Q9S7C0
45	48	49.5	1146	2	Q8P8T6

Q6va78	sars corona
Q6va89	sars corona
Q6vaa0	sars corona
Q6wgp3	sars corona
P83692	thielavia h
Q8cmx4	methanosarc
Q6ty54	candida gla
Q8yv57	anabaena sp
Q6fp15	candida gla
Q6udw6	plasmodium
Q9ca95	arabidopsis
Q8vz83	arabidopsis
Q9s7c0	arabidopsis
Q8p8t6	xanthomonas

ALIGNMENTS

RESULT 1

VGL2_CVHSA STANDARD: PRT: 1255 AA.
AC P59594, Q7T696; Q7TA19, Q7TFA2; Q7TFB1; Q80BV6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DB E2 glycoprotein precursor (Spike glycoprotein) (Peplomer protein).
CN Name=S;
OS Human coronavirus (strain SARS) (HCoV-SARS) (SARS-CoV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=227859;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate Urbani;
RX MEDLINE=22660724; PubMed=12730500; DOI=10.1126/science.1085952;
RA Rota P.A., Oberste M.S., Monroe S.S., Nix W.A., Campagnoli R.,
RA Icenogle J.P., Penaranda S., Bankamp B., Maher K., Chen M.-H.,
RA Tong S., Tamin A., Lowe L., Frace M., DeRisi J.L., Chen Q., Wang D.,
RA Erdman D.D., Peret T.C.T., Burns C., Ksiazek T.G., Rollin P.E.,
RA Sanchez A., Liffick S., Holloway B., Limor J., McCausland K.,
RA Olsen-Rasmussen M., Fouchier R., Guenther S., Osterhaus A.D.M.E.,
RA Drosten C., Pallansch M.A., Anderson L.J., Bellini W.J.;
RA "Characterization of a novel coronavirus associated with severe acute
RA respiratory syndrome.";
RL Science 300:1394-1399(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate Tor2;
RX MEDLINE=22660725; PubMed=12730501; DOI=10.1126/science.1085953;
RA Marra M.A., Jones S.J.M., Astell C.R., Holt R.A., Brooks-Wilson A.,
RA Butterfield Y.S.N., Khattar J., Asano J.K., Barber S.A., Chan S.Y.,
RA Cloutier A., Coughlin S.M., Freeman D., Girn N., Griffith O.L.,
RA Leach S.R., Mayo M., McDonald H., Montgomery S.B., Pandolf P.K.,
RA Petrescu A.S., Robertson A.G., Schein J.E., Siddiqui A., Smillius D.E.,
RA Stott J.M., Yang G.S., Plummer F., Anderson A., Artsob H., Bastien N.,
RA Bernard K., Booth T.F., Bowmass D., Czub M., Drebot M., Fernando L.,
RA Flick R., Garbutt M., Gray M., Grolla A., Jones S., Feldmann H.,
RA Meyers A., Kabani A., Li Y., Norman S., Stroher U., Tipples G.A.,
RA Tyler S., Vogrig R., Ward D., Watson B., Brunham R.C., Kraiden M.,
RA Petric M., Skowronski D.M., Upton C., Roper R.L.;
RA "The genome sequence of the SARS-associated coronavirus.";
RL Science 300:1399-1404(2003).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate CUHK-Su10, and Isolate CUHK-W1;
RX MEDLINE=22737955; PubMed=12853594; DOI=10.1056/NEJM200307103490216;
RA Tsui S.K.W., Chim S.S.C., Lo Y.M.D.;
RA "Coronavirus genomic-sequence variations and the epidemiology of the
RA severe acute respiratory syndrome.";
RL N. Engl. J. Med. 349:187-188(2003).
RN [4]
RP SEQUENCE FROM N.A.

RC STRAIN=Isolate HKU-39849;
 RX MEDLINE=22758472; PubMed=12876307;
 RA Zeng F.Y., Chan C.W., Chan M.N., Chen J.D., Chow K.Y.C., Hon C.C.C.,
 RA Hui R.K.H., Li J., Li V.Y.Y., Wang C.Y., Wang P.Y., Guan Y., Zheng B.,
 RA Poon L.L.M., Chan K.H., Yuen K.Y., Peiris J.S.M., Leung F.C.;
 RT "The complete genome sequence of severe acute respiratory syndrome
 RT coronavirus strain HKU-39849 (HK-39)";
 RL Exp. Biol. Med. 228:866-873(2003).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate G250, and Isolate HKU-36871;
 RX MEDLINE=22913660; PubMed=12958366; DOI=10.1126/science.1087139;
 RA Guan Y., Zheng B.J., He Y.Q., Liu X.L., Zhuang Z.X., Cheng C.L.,
 RA Luo S.W., Li P.H., Zhang L.J., Guan Y.J., Butt K.M., Wong K.L.,
 RA Chan K.W., Lim W., Shortridge K.F., Yuen K.Y., Peiris J.S.M.,
 RA Poon L.L.M.;
 RL "Isolation and characterization of viruses related to the SARS
 RT coronavirus from animals in southern China";
 RL Science 302:276-278(2003).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate BJ01, Isolate BJ02, Isolate BJ03, Isolate BJ04, and
 RC Isolate GD01;
 RA Qian E., Zhu Q., Yu M., Fan B., Chang G., Si B., Yang B., Peng W.,
 RA Jiang T., Liu B., Deng Y., Liu H., Zhang Y., Wang C., Li Y., Gan Y.,
 RA Li X., Lu F., Tan G., Yang R., Cao W.S., Wang J., Chen W., Cong L.,
 RA Deng Y., Dong W., Han Y., Hu W., Lei M., Li C., Li G., Li H.,
 RA Li S., Li S., Li W., Li W., Lin W., Liu J., Liu Z., Lu H., Ni P.,
 RA Qi Q., Sun Y., Tang L., Tong Z., Wang J., Wang X., Wu Q., Xi Y.,
 RA Xu Z., Yang L., Ye C., Ye J., Zhang B., Zhang F., Zhang J., Zhang X.,
 RA Zhou J., Yang H.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate Sin2500, Isolate Sin2677, Isolate Sin2679,
 RC Isolate Sin2748, and Isolate sin2774;
 RX MEDLINE=22667074; PubMed=12781537; DOI=10.1016/S0140-6736(03)13414-9;
 RA Ruan Y., Wei C.L., Ling A.E., Vega V.B., Thoreau H., Se Thoe S.Y.,
 RA Chia J.-M., Ng P., Chiu K.P., Lim L., Zhang T., Chan K.P., Oon L.E.L.,
 RA Ng M.L., Leo S.Y., Ng L.F.P., Ren E.C., Stanton L.W., Long P.M.,
 RA Liu E.T.;
 RL "Comparative full-length genome sequence analysis of 14 SARS
 RT coronavirus isolates and common mutations associated with putative
 RT origins of infection";
 RL Lancet 361:1779-1785(2003).
 RN [8]
 RP ERRATUM.
 RX PubMed=12781581;
 RA Ruan Y., Wei C.L., Ling A.E., Vega V.B., Thoreau H., Se Thoe S.Y.,
 RA Chia J.-M., Ng P., Chiu K.P., Lim L., Zhang T., Chan K.P., Oon L.E.L.,
 RA Ng M.L., Leo S.Y., Ng L.F.P., Ren E.C., Stanton L.W., Long P.M.,
 RA Liu E.T.;
 RL Lancet 361:1832-1832(2003).
 RN [9]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate TW1;
 RA Yeh S.-H., Kao C.-L., Tsai C.-Y., Liu C.-J., Chen D.-S., Chen P.-J.;
 RT "The complete genome of SARS coronavirus clone TW1";
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 RN [10]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate FRA;
 RA Eichmann M., Becker S., Klenk H.-D., Doerr H.W., Stadler K.,
 RA Censini S., Guidotti S., Maignani V., Scarselli M., Mora M.,
 RA Donati C., Han J., Song H.C., Abrignani S., Covacci A., Rappuoli R.;
 RT "SARS virus is a close relative of type II Coronaviruses";
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 RN [11]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate Frankfurt 1;
 RA Thiel V., Hertzog T., Putics A., Ivanov K.A., Schelle B., Bayer S.,
 RA Scheiner B., Weiland H., Weissbrich B., Ziebuhr J.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 RN [12]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate TWC;
 RA Yang J.-Y., Lin J.-H., Chiu S.-C., Wang S.-F., Lee S.-C., Lin Y.-C.,
 RA Hsu C.-K., Chen H.-Y., Chang J.G., Chen P.-J., Su I.-J.;
 RT "Genomic sequence of SARS isolate from the first fatal case in
 RT Taiwan";
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 RN [13]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate ZJ01;
 RA Cong L.-M., Ding G.-Q., Lu Y.-Y., Weng J.-Q., Yan J.-Y., Hu N.-P.,
 RA Wo J.-B., Chen S.-Y., Zhang Y.-J., Mei L.-L., Wang Z.-G., Yao J.,
 RA Zhu H.-P., Lu Q.-Y., Li M.-H., Gong L.-M., Shi W., Li L.-J.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 RN [14]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate Shanghai LY;
 RA Yuan Z., Zhang X., Hu Y., Lan S., Wang H., Zhou Z., Wen Y.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 RN [15]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate Taiwan TC1, Isolate Taiwan TC2, and Isolate Taiwan TC3;
 RA Chang J.-G.C., Lin T.-H., Chen C.-M., Lin C.-S., Chan W.-L.,
 RA Shih M.-C.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 RN [16]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate TWH, Isolate TWJ, Isolate TWK, Isolate TWS, and
 RC Isolate TWY;
 RA Shu H.-Y., Wu K.-M., Tsai S.-P.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 RN [17]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate HSR 1;
 RA Canducci F., Clementi M., Poli G., Vicenzi E.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 RN [18]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate TWC2, and Isolate TWC3;
 RA Yang J.-Y., Lin J.-H., Chiu S.-C., Wang S.-F., Lee H.-C., Lin Y.-C.,
 RA Hsu C.-K., Chen H.-Y., Chen P.-J., Su I.-J.;
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
 RN [19]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate AS;
 RA Balotta C., Corvasce S., Violin M., Galli M., Moroni M.,
 RA Vignani G.M., Ruan Y.J., Salemi M.;
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [20]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Isolate Shanghai QXC1;
 RA Yuan Z., Zhang X., Hu Y., Lan S., Wang H., Zhou Z., Wen Y.;
 RT "Analysis of SARS coronavirus genome in Shanghai isolates";
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 CC 1- FUNCTION: Structural protein that makes spikes at the surface of
 CC the virus. Determines enteropathogenicity and virulence of the
 CC virus (Potential).
 CC 1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC 1- DOMAIN: The spike S1 domain displays the specificity for the host
 CC receptor.
 CC 1- DOMAIN: The leucine zipper-like heptad repeats may mediate the
 CC fusion of viral and cellular membranes.
 CC 1- SIMILARITY: Contains 1 spike S1 domain.
 CC 1- SIMILARITY: Contains 1 spike S2 domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```
CC -----
DR EMBL; AY278741; AAP13441.1; -.

Query Match      100.0%; Score 97; DB 1; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGKPC 17
   |||||
Db 451 FERDISNVFSPDGKPC 467

RESULT 2
ID Q692E4 PRELIMINARY; PRT; 1255 AA.
AC Q692E4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Spike glycoprotein.
OS SARS coronavirus TJF.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=284672;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen W., Yan M., Liu M.;
RT "Isolation and Identification of Viruses Related to the SARS
   Coronavirus from swines in China.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY654624; AAT76147.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006944; P:membrane fusion; IEA.
DR GO; GO:0046813; P:virion attachment, binding of host cell sur. .; IEA.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01601; Corona_S2; 1.
DR SEQUENCE 1255 AA; 139153 MW; 508759C700289A0E CRC64;

Query Match      100.0%; Score 97; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGKPC 17
   |||||
Db 451 FERDISNVFSPDGKPC 467

RESULT 3
ID Q6DSU4 PRELIMINARY; PRT; 1255 AA.
AC Q6DSU4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Spike protein.
OS SARS coronavirus HHS-2004.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=285267;
RN [1]
RP SEQUENCE FROM N.A.
RA Song H., Ling N., Li Y., Zhu J., Wang E.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY648300; AAT74874.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006944; P:membrane fusion; IEA.
DR GO; GO:0046813; P:virion attachment, binding of host cell sur. .; IEA.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01601; Corona_S2; 1.
DR SEQUENCE 1255 AA; 139052 MW; 4587B501B7309C4D CRC64;

Query Match      100.0%; Score 97; DB 1; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGKPC 17
   |||||
Db 451 FERDISNVFSPDGKPC 467

RESULT 4
ID Q6GYR1 PRELIMINARY; PRT; 1255 AA.
AC Q6GYR1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE E2 glycoprotein.
OS SARS coronavirus LJJ-2004.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=273522;
RN [1]
RP SEQUENCE FROM N.A.
RA Li J., Wei W., Xiao W., Wang M., Wang J., Zhao J., Sun Z., Pei Y.,
   Chen Y.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY595412; AAT52330.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006944; P:membrane fusion; IEA.
DR GO; GO:0046813; P:virion attachment, binding of host cell sur. .; IEA.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01601; Corona_S2; 1.
DR SEQUENCE 1255 AA; 139170 MW; 90AF6E62485B525D CRC64;

Query Match      100.0%; Score 97; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGKPC 17
   |||||
Db 451 FERDISNVFSPDGKPC 467

RESULT 5
ID Q6JH38 PRELIMINARY; PRT; 1255 AA.
AC Q6JH38;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Spike protein.
OS SARS coronavirus Sino3-11.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=255729;
RN [1]
RP SEQUENCE FROM N.A.
RA Jin W.W., Feng J.D., Du Z.L., Hu L.X., Liu Y.X., Liu Y.F., Zhang X.M.,
   Gao H., Ning Y., Zhang J.S., Li N., Yin W.D.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY485278; AAR23258.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006944; P:membrane fusion; IEA.
DR GO; GO:0046813; P:virion attachment, binding of host cell sur. .; IEA.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01601; Corona_S2; 1.
DR SEQUENCE 1255 AA; 139124 MW; 1C49ACA2CFD38FC0 CRC64;
```

```
Query Match      100.0%; Score 97; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGKPC 17
   |||||
Db 451 FERDISNVFSPDGKPC 467

RESULT 4
ID Q6GYR1 PRELIMINARY; PRT; 1255 AA.
AC Q6GYR1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE E2 glycoprotein.
OS SARS coronavirus LJJ-2004.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=273522;
RN [1]
RP SEQUENCE FROM N.A.
RA Li J., Wei W., Xiao W., Wang M., Wang J., Zhao J., Sun Z., Pei Y.,
   Chen Y.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY595412; AAT52330.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006944; P:membrane fusion; IEA.
DR GO; GO:0046813; P:virion attachment, binding of host cell sur. .; IEA.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01601; Corona_S2; 1.
DR SEQUENCE 1255 AA; 139170 MW; 90AF6E62485B525D CRC64;

Query Match      100.0%; Score 97; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGKPC 17
   |||||
Db 451 FERDISNVFSPDGKPC 467

RESULT 5
ID Q6JH38 PRELIMINARY; PRT; 1255 AA.
AC Q6JH38;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Spike protein.
OS SARS coronavirus Sino3-11.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=255729;
RN [1]
RP SEQUENCE FROM N.A.
RA Jin W.W., Feng J.D., Du Z.L., Hu L.X., Liu Y.X., Liu Y.F., Zhang X.M.,
   Gao H., Ning Y., Zhang J.S., Li N., Yin W.D.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY485278; AAR23258.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006944; P:membrane fusion; IEA.
DR GO; GO:0046813; P:virion attachment, binding of host cell sur. .; IEA.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01601; Corona_S2; 1.
DR SEQUENCE 1255 AA; 139124 MW; 1C49ACA2CFD38FC0 CRC64;
```

Query Match 100.0%; Score 97; DB 2; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 1.7e-06;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FERDISNVFSPDGKPC 17
 |||||
 Db 451 FERDISNVFSPDGKPC 467

RESULT 6

Q6JH46 Q6JH46 PRELIMINARY; PRT; 1255 AA.
 AC Q6JH46;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Spike protein.
 OS SARS coronavirus Sinol-11.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=255730;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sinol-11;
 RA Jin W., Feng J.D., Du Z.L., Hu L.X., Liu Y.X., Liu Y.F., Zhang X.M.,
 RA Jin W., Ning Y., Zhang J.S., Li N., Yin W.D.;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY485277; AAR3250.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0006944; P:membrane fusion; IEA.
 DR GO; GO:0046813; P:virion attachment, binding of host cell sur. . .; IEA.
 DR InterPro; IPR002552; Corona_S2.
 DR Pfam; PF01601; Corona_S2; 1.
 DR PIR; P01601; Corona_S2; 1.
 SQ SEQUENCE 1255 AA; 139048 MW; 80E6F04D325D3E00 CRC64;

Query Match 100.0%; Score 97; DB 2; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 1.7e-06;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FERDISNVFSPDGKPC 17
 |||||
 Db 451 FERDISNVFSPDGKPC 467

RESULT 7

Q6QU82 Q6QU82 PRELIMINARY; PRT; 1255 AA.
 AC Q6QU82;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Spike glycoprotein.
 OS SARS coronavirus GD03T0013.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=262338;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zheng K., Zhou H.Q., Chen Q.X., Zhang X., Zheng H.Y., Wan Z.Y.;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY525636; AAS10463.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0006944; P:membrane fusion; IEA.
 DR GO; GO:0046813; P:virion attachment, binding of host cell sur. . .; IEA.
 DR InterPro; IPR002552; Corona_S2.
 DR Pfam; PF01601; Corona_S2; 1.
 SQ SEQUENCE 1255 AA; 139173 MW; DD62DB871695C3A5 CRC64;

Query Match 100.0%; Score 97; DB 2; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 1.7e-06;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FERDISNVFSPDGKPC 17
 |||||
 Db 451 FERDISNVFSPDGKPC 467

RESULT 8

Q6R7Y6 Q6R7Y6 PRELIMINARY; PRT; 1255 AA.
 AC Q6R7Y6;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Spike glycoprotein S.
 OS SARS coronavirus NS-1.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=260743;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NS-1;
 RA Zheng H., Wei C., Ding R., Wang W., Li W., Wang J., Tao W., Yu X.,
 RA Guo H., Chen J., Wei W., Li J., Zhang Y., Wang X., Sun Y., Jiao J.,
 RA Wang Y., Zhou C.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY508724; AAR91586.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0006944; P:membrane fusion; IEA.
 DR GO; GO:0046813; P:virion attachment, binding of host cell sur. . .; IEA.
 DR InterPro; IPR002552; Corona_S2.
 DR Pfam; PF01601; Corona_S2; 1.
 DR PIR; P01601; Corona_S2; 1.
 SQ SEQUENCE 1255 AA; 139087 MW; 7C49A690C54F25C1 CRC64;

Query Match 100.0%; Score 97; DB 2; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 1.7e-06;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FERDISNVFSPDGKPC 17
 |||||
 Db 451 FERDISNVFSPDGKPC 467

RESULT 9

Q6RCW5 Q6RCW5 PRELIMINARY; PRT; 1255 AA.
 AC Q6RCW5;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Putative spike glycoprotein.
 OS SARS coronavirus TW5.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=258972;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14983045; DOI=10.1073/pnas.0307904100;
 RG National Taiwan University SARS Research Team;
 RA Yeh S.-H., Wang H.-Y., Tsai C.-Y., Kao C.-L., Yang J.-Y., Liu H.-W.,
 RA Su I.-J., Tsai S.-F., Chen D.-S., Chen P.-J.;
 RT "Characterization of severe acute respiratory syndrome coronavirus
 genomes in Taiwan: molecular epidemiology and genome evolution.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:2542-2547(2004).
 DR EMBL; AY502932; AAR87600.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0006944; P:membrane fusion; IEA.
 DR GO; GO:0046813; P:virion attachment, binding of host cell sur. . .; IEA.
 DR InterPro; IPR002552; Corona_S2.
 DR Pfam; PF01601; Corona_S2; 1.
 DR PIR; P01601; Corona_S2; 1.
 SQ SEQUENCE 1255 AA; 139124 MW; 1C49ACA2CFD38FC0 CRC64;


```
Query Match 100.0%; Score 97; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGKPC 17
Db 451 FERDISNVFSPDGKPC 467

RESULT 10
Q6RCX6 PRELIMINARY; PRT; 1255 AA.
AC Q6RCX6;
AC Q6RCX6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative spike glycoprotein.
OS SARS coronavirus TW5.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=258971;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14983045; DOI=10.1073/pnas.0307904100;
RG National Taiwan University SARS Research Team;
RA Yeh S.-H., Wang H.-Y., Tsai C.-Y., Kao C.-L., Yang J.-Y., Liu H.-W.,
RA Su I.-J., Tsai S.-F., Chen D.-S., Chen P.-J.;
RT "Characterization of severe acute respiratory syndrome coronavirus
RT genomes in Taiwan: molecular epidemiology and genome evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2542-2547(2004).
DR EMBL; AY502931; AAR87589.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006944; P:membrane fusion; IEA.
DR GO; GO:0046813; P:virion attachment, binding of host cell sur. . .; IEA.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01601; Corona_S2; 1.
DR SEQUENCE 1255 AA; 139124 MW; 1C49ACA2CFD38FC0 CRC64;

Query Match 100.0%; Score 97; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGKPC 17
Db 451 FERDISNVFSPDGKPC 467

RESULT 11
Q6RCY7 PRELIMINARY; PRT; 1255 AA.
AC Q6RCY7;
AC Q6RCY7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative spike glycoprotein.
OS SARS coronavirus TW7.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=258970;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14983045; DOI=10.1073/pnas.0307904100;
RG National Taiwan University SARS Research Team;
RA Yeh S.-H., Wang H.-Y., Tsai C.-Y., Kao C.-L., Yang J.-Y., Liu H.-W.,
RA Su I.-J., Tsai S.-F., Chen D.-S., Chen P.-J.;
RT "Characterization of severe acute respiratory syndrome coronavirus
RT genomes in Taiwan: molecular epidemiology and genome evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2542-2547(2004).
DR EMBL; AY502930; AAR87578.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006944; P:membrane fusion; IEA.
```

```
DR GO; GO:0046813; P:virion attachment, binding of host cell sur. . .; IEA.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01601; Corona_S2; 1.
DR SEQUENCE 1255 AA; 139124 MW; 1C49ACA2CFD38FC0 CRC64;

Query Match 100.0%; Score 97; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGKPC 17
Db 451 FERDISNVFSPDGKPC 467

RESULT 12
Q6RCZ8 PRELIMINARY; PRT; 1255 AA.
AC Q6RCZ8;
AC Q6RCZ8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative spike glycoprotein.
OS SARS coronavirus TW6.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=258969;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14983045; DOI=10.1073/pnas.0307904100;
RG National Taiwan University SARS Research Team;
RA Yeh S.-H., Wang H.-Y., Tsai C.-Y., Kao C.-L., Yang J.-Y., Liu H.-W.,
RA Su I.-J., Tsai S.-F., Chen D.-S., Chen P.-J.;
RT "Characterization of severe acute respiratory syndrome coronavirus
RT genomes in Taiwan: molecular epidemiology and genome evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2542-2547(2004).
DR EMBL; AY502929; AAR87567.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006944; P:membrane fusion; IEA.
DR GO; GO:0046813; P:virion attachment, binding of host cell sur. . .; IEA.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01601; Corona_S2; 1.
DR SEQUENCE 1255 AA; 139124 MW; 1C49ACA2CFD38FC0 CRC64;

Query Match 100.0%; Score 97; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FERDISNVFSPDGKPC 17
Db 451 FERDISNVFSPDGKPC 467

RESULT 13
Q6RD09 PRELIMINARY; PRT; 1255 AA.
AC Q6RD09;
AC Q6RD09;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative spike glycoprotein.
OS SARS coronavirus TW5.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=258968;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14983045; DOI=10.1073/pnas.0307904100;
RG National Taiwan University SARS Research Team;
RA Yeh S.-H., Wang H.-Y., Tsai C.-Y., Kao C.-L., Yang J.-Y., Liu H.-W.,
RA Su I.-J., Tsai S.-F., Chen D.-S., Chen P.-J.;
RT "Characterization of severe acute respiratory syndrome coronavirus
RT genomes in Taiwan: molecular epidemiology and genome evolution.";
```

```
RL Proc. Natl. Acad. Sci. U.S.A. 101:2542-2547(2004).
DR EMBL; AY502928; AAR87556.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006944; P:membrane fusion; IEA.
DR GO; GO:0046813; P:virion attachment, binding of host cell sur. . .; IEA.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01601; Corona_S2; 1.
SQ SEQUENCE 1255 AA; 139124 MW; 1C49ACA2CFD38FC0 CRC64;

Query Match 100.0%; Score 97; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FERDISNVFSPDGKPC 17
Db 451 FERDISNVFSPDGKPC 467

RESULT 14
Q6RD20 PRELIMINARY; PRT; 1255 AA.
AC Q6RD20;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative spike glycoprotein.
OS SARS coronavirus TW4.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=258967;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14983045; DOI=10.1073/pnas.0307904100;
RG National Taiwan University SARS Research Team;
RA Yeh S.-H., Wang H.-Y., Tsai C.-Y., Kao C.-L., Yang J.-Y., Liu H.-W.,
RA Su I.-J., Tsai S.-F., Chen D.-S., Chen P.-J.;
RT "Characterization of severe acute respiratory syndrome coronavirus
genomes in Taiwan: molecular epidemiology and genome evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2542-2547(2004).
DR EMBL; AY502927; AAR87545.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006944; P:membrane fusion; IEA.
DR GO; GO:0046813; P:virion attachment, binding of host cell sur. . .; IEA.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01601; Corona_S2; 1.
SQ SEQUENCE 1255 AA; 139124 MW; 1C49ACA2CFD38FC0 CRC64;

Query Match 100.0%; Score 97; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FERDISNVFSPDGKPC 17
Db 451 FERDISNVFSPDGKPC 467

RESULT 15
Q6RD31 PRELIMINARY; PRT; 1255 AA.
ID Q6RD31
AC Q6RD31;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative spike glycoprotein.
OS SARS coronavirus TW3.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=258966;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14983045; DOI=10.1073/pnas.0307904100;
```

```
RG National Taiwan University SARS Research Team;
RA Yeh S.-H., Wang H.-Y., Tsai C.-Y., Kao C.-L., Yang J.-Y., Liu H.-W.,
RA Su I.-J., Tsai S.-F., Chen D.-S., Chen P.-J.;
RT "Characterization of severe acute respiratory syndrome coronavirus
genomes in Taiwan: molecular epidemiology and genome evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:2542-2547(2004).
DR EMBL; AY502926; AAR87534.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0006944; P:membrane fusion; IEA.
DR GO; GO:0046813; P:virion attachment, binding of host cell sur. . .; IEA.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01601; Corona_S2; 1.
SQ SEQUENCE 1255 AA; 139124 MW; 1C49ACA2CFD38FC0 CRC64;

Query Match 100.0%; Score 97; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FERDISNVFSPDGKPC 17
Db 451 FERDISNVFSPDGKPC 467

Search completed: March 18, 2005, 19:06:08
Job time : 157.72 secs
```